

Generating maximally disassortative graphs with given degree distribution.

Pim van der Hoorn¹, Liudmila Ostroumova Prokhorenkova^{2,3} and Egor Samosvat^{2,3}

¹University of Twente, Enschede, the Netherlands

²Moscow Institute of Physics and Technology, Moscow, Russia

³Yandex, Moscow, Russia

July 7, 2016

Abstract

In this paper we consider the optimization problem of generating graphs with a prescribed degree distribution, such that the correlation between the degrees of connected nodes, as measured by Spearman's rho, is minimal. We provide an algorithm for solving this problem and obtain a complete characterization of the joint degree distribution in these maximally disassortative graphs, in terms of the size-biased degree distribution. As a result we get a lower bound for Spearman's rho on graphs with an arbitrary given degree distribution. We use this lower bound to show that for any fixed tail exponent, there exist scale-free degree sequences with this exponent such that the minimum value of Spearman's rho for all graphs with such degree sequences is arbitrary close to zero. This implies that specifying only the tail behavior of the degree distribution, as is often done in the analysis of complex networks, gives no guarantees for the minimum value of Spearman's rho.

Keywords: graphs, degree distribution, degree-degree correlation, disassortativity, scale-free distribution

1 Introduction

An important second order characteristic of the topology of a graph, introduced in [12], is the correlation between the degrees at both sides of a randomly sampled edge, also called degree-degree correlation or degree assortativity. A graph is called assortative, or is said to have assortative mixing, if this correlation is positive and disassortative if it is negative. In assortative graphs, nodes of a certain degree have a preference to connect to nodes of similar degree, while in a disassortative graph the opposite is true, for instance, nodes of small degrees connect to nodes with large degrees. When the degrees of connected nodes are uncorrelated the graph is said to have neutral mixing.

Recently, the problem of generating graphs with a given joint degree structure has been investigated. In [2] and [14] algorithms are introduced for constructing and sampling graphs with a given joint degree matrix J , where an entry $J_{k\ell}$ denotes the number of edges between nodes of degrees k and ℓ . An algorithm for generating random graphs whose joint degree distribution converges to a given limiting distribution is given in [5] and [6] under the assumption that the degrees are uniformly bounded in the size of the graph.

A different branch of research is concerned with generating graphs that have extreme degree-degree correlation structure, either maximally assortative or disassortative, and analyzing structural properties of such graphs. One algorithm that is often used for this is the so-called edge swap algorithm [7, 8, 20]. In the context of degree-degree correlations, this algorithm starts from

an initial graph, with a prescribed degree sequence, and in each step two edges are sampled and switched based on some rule, in order to obtain a maximally (dis)assortative graph. In [9] this algorithm is used to obtain scaling results for Pearson's correlation coefficient, as introduced in [12] on maximally (dis)assortative graphs where the degrees follow a scale-free distribution. The results from [9] are extended in [21], where a lower bound for Pearson's correlation coefficient is established in scale-free graphs.

One of the problems with the current analysis of graphs with extreme degree-degree correlation structure is the use of Pearson's correlation coefficient as a measure for assortativity, since this measure has been shown to be size-dependent when the degree distribution has infinite variance [16, 18]. In these papers new, rank-based, correlation measures are introduced and it is shown that these measures converge to a proper limit, determined by the joint degree distribution, under very standard assumptions, see [16, 17]. Therefore, in this paper, we follow their suggestion and use a rank correlation measure related to Spearman's rho.

We introduce a greedy algorithm for generating graphs with a given degree distribution that are maximally disassortative, with respect to the rank correlation measure Spearman's rho. The algorithm gives insights into the joint degree structure of these graphs. Using these insights we are able to characterize the limiting joint degree distribution of maximally disassortative graphs, in terms of the size-biased degree distribution. Moreover, due to use of a general framework describing the convergence of the empirical distributions, we are able to characterize the speed of the convergence.

An important consequence of the joint degree structure of maximally disassortative graphs is that the tail of the distribution does not affect the minimum value of Spearman's rho. Moreover, we are able to construct regularly varying distributions with a prescribed exponent, such that Spearman's rho on any graph with this degree distribution is bounded from below by a value that is arbitrary close to zero.

We complement our theoretical results with simulations that show the concentration of Spearman's rho for graphs generated by our algorithm and illustrate how this measure is influenced by the shape of the size-biased degree distribution. We observe that the minimal value Spearman's rho becomes larger when more mass is placed in the head of the degree distribution, while increasing the mass in the tail of the distribution decreases this value.

2 Notations and results

We will start by introducing some notation and summarizing our main results.

2.1 Graphs and Degree sequences

Given a degree sequence $\mathbf{D}_n = \{D_1, D_2, \dots, D_n\}$ we define $L_n = \sum_{i=1}^n D_i$. That is, L_n is the sum of the degrees and hence *twice* the number of edges in a graph with degree sequence \mathbf{D}_n . We further define the empirical and sized-biased degree distributions by, respectively,

$$f_n(k) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}, \quad (1)$$

$$f_n^*(k) = \frac{1}{L_n} \sum_{i=1}^n k \mathbb{1}_{\{D_i=k\}}, \quad (2)$$

and let F_n and F_n^* be the corresponding cumulative distribution functions.

We will assume that the empirical distributions f_n and f_n^* converge to certain limiting distributions f and f^* as follows.

Assumption 2.1. *Let f and f^* be probability mass functions on the non-negative integers such that*

$$\sum_{k=0}^{\infty} k^{1+\eta} f(k) < \infty \quad (3)$$

for some $\eta > 0$ and if we define, for some $\varepsilon > 0$,

$$\Omega_n = \left\{ \max \left\{ \sum_{k=0}^{\infty} \left| \sum_{t=0}^k f_n(t) - f(t) \right|, \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \right\} \leq n^{-\varepsilon} \right\},$$

then

$$\mathbb{P}(\Omega_n) \xrightarrow{\mathbb{P}} 1 \quad \text{as } n \rightarrow \infty.$$

We will denote by F and F^* the cumulative distributions of f and f^* , respectively. Since we assume that the event Ω_n occurs, asymptotically, with probability one, we will often use the probability of its complement Ω_n^c to describe the speed of convergence in our results. In addition, for simplicity of notation, we will assume throughout this paper that $f(k), f^*(k) > 0$ for all $k \geq 0$. Our results extend in a straightforward manner to other cases, by considering only all k for which $f(k), f^*(k) > 0$.

To give some explanation regarding Assumption 2.1 we remark that the first expression in the maximum of the event Ω_n is related to the Kantorovich-Rubinstein distance or, equivalently, the Wasserstein metric of order one between the distributions F_n and F . Convergence in this metric is equivalent to weak convergence as well as convergence of the first absolute moments, see [19] for more details. Hence, assumption 2.1 describes the joint convergence of f_n to f and f_n^* to f^* in the Kantorovich-Rubinstein distance and the 1-norm, respectively. We used different metrics for the convergence of f_n and f_n^* , since the Wasserstein metric is only a true distance when the distributions have finite first absolute moment. We are not assuming that the distribution f^* has finite first absolute moment since we want to consider graphs whose degree distributions have infinite second moment, which implies that the size-biased degree distribution has infinite mean.

In order to state our results we will use the following definition

Definition 2.2. Let \mathbf{D}_n be a degree sequence. We say that $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ if and only if \mathbf{D}_n satisfies Assumption 2.1 with density functions f and f^* and $\eta, \varepsilon > 0$. For a graph G_n with a degree sequence \mathbf{D}_n , we will write $G_n \in \mathcal{G}_{\eta, \varepsilon}(f, f^*)$ if $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$.

2.2 Spearman's rho on graphs

For an integer valued random variable X , we denote its cumulative distribution function by F_X and define

$$\mathcal{F}_X(k) = F_X(k) + F_X(k-1), \quad \text{for all } k \in \mathbb{Z}. \quad (4)$$

Now, let X and Y be two integer valued random variables, then Spearman's rho is defined as [10]

$$\rho(X, Y) = 3\mathbb{E}[\mathcal{F}_X(X)\mathcal{F}_Y(Y)] - 3. \quad (5)$$

For the definition of Spearman's rho on graphs it is convenient to consider directed edges. To make this work on undirected graphs we replace each edge $i - j$ by two edges, $i \rightarrow j$ and $j \rightarrow i$. We refer to this graph as the bi-directed version of the original graph. Although the graph on which Spearman's rho is computed is directed, we will not distinguish between this and the original undirected graph G_n . That is, we will write $i \rightarrow j \in G_n$ to mean that $i \rightarrow j$ is present in the bi-directed version of G_n , which is equivalent to stating that $i - j \in G_n$. We recall that L_n denotes the sum over all degrees, so that L_n is *twice* the number of undirected edges and equal to the corresponding number of directed edges in G_n .

Next we will consider Spearman's rho with uniform ranking, as described in [16] and [18]. That is, we take $(\mathbf{U}_{i \rightarrow j}, \mathbf{W}_{i \rightarrow j})$ to be a vector of independent uniform random variables $U_{i \rightarrow j}$ and $W_{i \rightarrow j}$ on $(0, 1)$, for each edge $i \rightarrow j \in G$, and define the ranking functions $R_*(i \rightarrow j)$ and $R^*(i \rightarrow j)$ by

$$\begin{aligned} R_*(i \rightarrow j) &= \sum_{s \rightarrow t \in G} \mathbb{1}_{\{D_s + U_{s \rightarrow t} \geq D_i + U_{i \rightarrow j}\}}, \\ R^*(i \rightarrow j) &= \sum_{s \rightarrow t \in G} \mathbb{1}_{\{D_t + W_{s \rightarrow t} \geq D_j + W_{i \rightarrow j}\}}, \end{aligned}$$

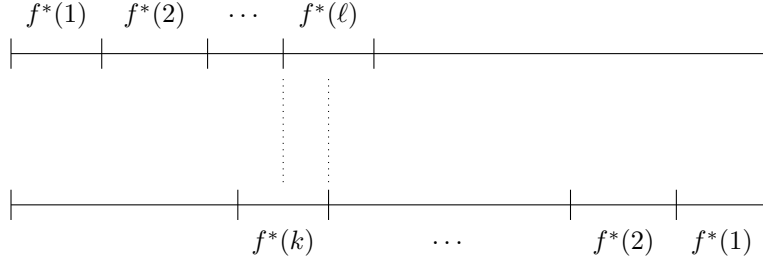


Figure 1: Illustration of the functions ψ and \mathcal{E} .

where we let $\sum_{i \rightarrow j \in G}$ denote the sum over all edges $i \rightarrow j$ in the graph G . With these definitions, Spearman's rho is defined as, see [16, 18],

$$\rho(G_n) = \frac{12 \sum_{i \rightarrow j \in G} R_*(i \rightarrow j) R^*(i \rightarrow j) - 3L_n(L_n + 1)^2}{L_n^3 - L_n}. \quad (6)$$

To link $\rho(G_n)$ to (5), let h_n denote the empirical joint probability density function of the degrees on both sides of a random edge, i.e.

$$h_n(k, \ell) = \frac{1}{L_n} \sum_{i \rightarrow j \in G} \mathbb{1}_{\{D_i=k\}} \mathbb{1}_{\{D_j=\ell\}}.$$

Then, if h_n converges to some limiting distribution h , it follows from Theorem 3.2 in [16] that

$$\rho(G_n) \xrightarrow{\mathbb{P}} \rho(X, Y) \quad \text{as } n \rightarrow \infty,$$

where X and Y have joint distribution h . In other words, $\rho(G_n)$ is a consistent estimator of $\rho(X, Y)$. Moreover, in [18] it is shown that $\rho(G_n)$ is asymptotically equivalent to

$$\tilde{\rho}(G_n) = \frac{3}{L_n} \sum_{i \rightarrow j} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j) - 3. \quad (7)$$

Since this expression is easier to analyze mathematically, we will use this measures in our statements. We show with numerical experiments in Section 5 that our results also hold for the original expression 6.

2.3 Main results

In order to state the first result we define, for any $k, \ell \geq 1$, the functions

$$\psi(k, \ell) = \mathbb{1}_{\{1 - F^*(k) < F^*(\ell)\}} \mathbb{1}_{\{1 - F^*(k-1) > F^*(\ell-1)\}}, \quad (8)$$

$$\mathcal{E}(k, \ell) = \min(1 - F^*(k-1), F^*(\ell)) - \max(1 - F^*(k), F^*(\ell-1)). \quad (9)$$

These functions can be understood as follows. Consider the partition of the interval $[0, 1]$ given by the sequence $\{f^*(1), f^*(2), \dots\}$. Now take a copy of this partitioned interval, reverse it and place it below the original interval, see Figure 1. Then $\psi(k, \ell)$ is the indicator of the event that the interval corresponding to $f^*(\ell)$ on the top intersects with the interval corresponding to $f^*(k)$ at the bottom, while $\mathcal{E}(k, \ell)$ is the size of this intersection.

With these functions we now define the joint probability density function

$$h(k, \ell) = \psi(k, \ell) \mathcal{E}(k, \ell), \quad k, \ell = 1, 2, \dots \quad (10)$$

Our main result states that if X and Y have joint distribution h , then Spearman's rho on graphs with a degree sequence satisfying Assumption 2.1 is bounded from below by $\rho(X, Y)$, and that this minimum is attained for a specific sequence of graphs.

Theorem 2.1. Let $G_n \in \mathcal{G}_{\eta, \varepsilon}(f, f^*)$ and let X, Y be random variables with joint distribution h as defined in (10). Then, for any $0 < \delta < \min(\varepsilon, 1/2)$ and $K > 0$,

$$\mathbb{P}(\tilde{\rho}(G_n) \geq \tilde{\rho}(D_*, D^*) - Kn^{-\delta}) \geq 1 - O(n^{-\varepsilon+\kappa} + \mathbb{P}(\Omega_n^c)), \quad (11)$$

as $n \rightarrow \infty$, where

$$\kappa = \frac{\varepsilon + \delta}{2}.$$

Moreover, there exists graphs \hat{G}_n with the same degree sequence as G_n , such that, as $n \rightarrow \infty$,

$$\mathbb{P}\left(\left|\tilde{\rho}(\hat{G}_n) - \rho(D_*, D^*)\right| > Kn^{-\delta}\right) \leq O(n^{-\varepsilon+\kappa} + \mathbb{P}(\Omega_n^c)).$$

This result can be understood in terms of the following optimization problem. Given a degree sequences $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$, define

$$\mathcal{F}_n^*(k) = F_n^*(k) + F_n^*(k-1).$$

and consider, for fixed n , the following objective function

$$\min_{G \in \mathcal{G}(\mathbf{D}_n)(f, f^*)} \frac{1}{L_n} \sum_{i \rightarrow j \in G} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j), \quad (12)$$

where the minimum is understood to be taken over all graphs G_n with degree sequences satisfying Assumption 2.1 with density functions f and f^* . Then Theorem 2.1 states that with high probability

$$\min_{G_n \in \mathcal{G}_{\eta, \varepsilon}(f, f^*)} \rho(G_n) = \rho(D_*, D^*),$$

where $\rho(D_*, D^*)$ is given by, see (5),

$$\rho(D_*, D^*) = \sum_{k, \ell=0}^{\infty} \mathcal{F}^*(k) \mathcal{F}^*(\ell) \psi(k, \ell) \mathcal{E}(k, \ell),$$

with ψ and \mathcal{E} as defined in (8) and (9), respectively. Moreover, Theorem 2.1 provides a sequence of graphs \hat{G}_n that attains this minimum, i.e. a sequence of maximally disassortative graphs with the degree distribution f . These graphs will be generated by our algorithm, which we will present in Section 3.

We remark that although Theorem 2.1 solves the minimization problem of degree-degree correlations in undirected graphs by giving an the asymptotic minimum $\rho(D_*, D^*)$ on Spearman's rho, this minimum is, in general, hard to derive since it depends on the full size-biased limit density f^* . However, for specific cases it can be computed numerically by computing

$$\sum_{k=0}^K \sum_{\ell=0}^L \mathcal{F}^*(k) \mathcal{F}^*(\ell) \psi(k, \ell) \mathcal{E}(k, \ell),$$

for certain upper bounds K and L .

Part of the proof of Theorem 2.1 consists of showing that h is the limiting joint degree distribution of maximally disassortative graphs. From the interpretation of the functions ψ and \mathcal{E} , it follows that for all $k \geq K$, for some threshold K , all intervals corresponding to $f^*(k)$ on the top will be contained in the interval $f^*(1)$ at the bottom and vice versa. This implies that the large degree nodes will, asymptotically, all be connected to nodes with degree one. As a consequence we have that the tail of the distribution f^* , and hence also that of f , plays a negligible role in the lower bound of Spearman's rho. Therefore, we can construct degree distributions with specified tail behavior so that Spearman's rho on such graphs approaches zero from below, with arbitrary precision.

Theorem 2.2. *Let f be any probability density function with support on the non-negative integers, mean μ and*

$$\sum_{k=0}^{\infty} k^{1+\eta} f(k) < \infty,$$

for some $\eta > 0$. Then, for any $-1 < \rho < 0$, there exists a probability density function f_ρ on the non-negative integers with mean μ_ρ , which satisfies

$$\lim_{k \rightarrow \infty} \frac{1 - F_\rho(k)}{1 - F(k)} = \frac{\mu_\rho}{\mu}.$$

Moreover, for any sequence of graphs $G_n \in \mathcal{G}_{\eta, \varepsilon}(f_\rho, f_\rho^)$, where $f_\rho^*(k) = k f_\rho(k) / \mu_\rho$, we have*

$$\mathbb{P}(\tilde{\rho}(G_n) > \rho) \geq 1 - O\left(n^{-1+\kappa} + n^{-\varepsilon+3\kappa/4} + \mathbb{P}(\Omega_n^c)\right),$$

as $n \rightarrow \infty$, where $\kappa = \min(\varepsilon, 1/2)$.

The main message of Theorem 2.2 is that it is not the tail of the degree distribution that is crucial for the minimal value of $\tilde{\rho}(G_n)$.

The characterization of the tail of the degree distribution is most prominently present in the analysis of so-called scale-free networks. These are graphs where the limiting degree distribution F satisfies

$$1 - F(k) = \mathcal{L}(k)k^{-\gamma}, \quad \gamma > 1, \tag{13}$$

for some slowly varying function \mathcal{L} . The exponent γ is referred to as the tail exponent. As a corollary to Theorem 2.2 we obtain the following result which states that knowledge of only the tail exponent does not give any guarantees on the minimum value of Spearman's rho.

Corollary 2.3. *For any $-1 < \rho < 0$ and $\gamma > 1$, there exist distributions f and f^* , where F satisfies (13), such that for any sequence of graphs $G_n \in \mathcal{G}_{\eta, \varepsilon}(f, f^*)$*

$$\mathbb{P}(\rho(G_n) > \rho) \geq 1 - O\left(n^{-1+\kappa} + n^{-\varepsilon+3\kappa/4} + \mathbb{P}(\Omega_n^c)\right),$$

as $n \rightarrow \infty$, where $\kappa = \min(\varepsilon, 1/2)$.

2.4 Structure of the paper

The rest of this paper is structured as follows. In Section 3.1 we describe the algorithm for generating graphs that solves the optimization problem (12). A complete characterization of the empirical and limiting joint degree distribution is then given in Section 3.2. We describe the construction of degree sequences with arbitrary small value of Spearman's rho in Section 4. In Section 5 we illustrate our results by providing simulations for maximally disassortative graphs where the degrees follow a scale-free and a Poisson distribution. Finally, Section 6 contains all the proofs of our results.

3 Generating maximally disassortative graphs

We will describe an algorithm, called the **Disassortative Graph Algorithm (DGA)**, that solves (12).

3.1 The Disassortative Graph Algorithm

Any degree sequence \mathbf{D}_n can be represented by a list of stubs, where for each node i we have D_i stubs labeled i . A graph with degree sequence \mathbf{D}_n is then completely determined by the pairing

of the stubs. In order to describe our algorithm, let N_k denote the number of nodes with degree k and let z_n be the unique integer satisfying

$$\sum_{t=1}^{z_n} tN_t \geq \frac{L_n}{2} \quad \text{and} \quad \sum_{t=1}^{z_n-1} tN_t < \frac{L_n}{2}. \quad (14)$$

The idea of the **Disassortative Graph Algorithm** is to use z_n to divide the stubs in two columns. In the left column S_n we add the stubs belonging to nodes with high degree ($D_i \geq z_n$), in descending order. The right column T_n will be filled with stubs that belong to nodes with small degree ($D_i \leq z_n$) in ascending order. After this ordering we start pairing stubs from the left column to stubs in the right column, until we reach the first pair (i, j) for which $D_i = z_n = D_j$. We are now left with stubs belonging to nodes with degree z_n , hence the value of Spearman's rho (7) will not be influenced by the specific way in which we connect them. This means that we can, in principle, use any algorithm to connect these medium degree nodes. We will use the configuration model [3, 11, 13], more specifically the repeated configuration model, see Section 7.4 in [15]. The full algorithm is described in detail below.

Algorithm 1 Disassortative Graph Algorithm

- 1: Input: A degree sequence \mathbf{D}_n .
 - 2: Rank the nodes by their degrees in ascending order and let $\varrho(k)$ and denote the node with rank k , i.e. $D_{\phi(n)} \geq D_{\phi(n-1)} \geq \dots \geq D_{\phi(2)} \geq D_{\phi(1)}$.
 - 3: Create two empty lists S_n and T_n .
 - 4: Set $i = n$ and $j = 1$.
 - 5: **while** $D_{\phi(i)} \geq z_n$ **do**
 - 6: Fill the next $D_{\phi(i)}$ slots of S_n with stubs labeled $\phi(i)$.
 - 7: Set $i = i - 1$.
 - 8: **end while**
 - 9: **while** $D_{\phi(j)} \leq z_n$ **do**
 - 10: Add to T_n , $D_{\phi(j)}$ copies of stubs labeled: $\phi(j), \dots, \varrho(j + N_{D_{\phi(j)}} - 1)$.
 - 11: Set $j = j + N_{D_{\phi(j)}}$.
 - 12: **end while**
 - 13: Set $t = 1$, $i = S_n[t]$ and $j = T_n[t]$
 - 14: **while not** $D_i = z_n = D_j$ **do**
 - 15: Add edge $i - j$ to G_n .
 - 16: Set $t = t + 1$, $i = S_n[t]$ and $j = T_n[t]$.
 - 17: **end while**
 - 18: Set \mathbf{D}_n^z to be the degree sequence corresponding to the remaining unpaired stubs.
 - 19: Pair the stubs in \mathbf{D}_n^z using the configuration model.
 - 20: Output: G_n .
-

We will denote by G_n^* the induced sub-graph that has been created at the end of step 17 and let the complement $G_n^z = G_n \setminus G_n^*$ denote the graph generated by the configuration model in step 19. In addition we will write $G_n = \text{DGA}(\mathbf{D}_n)$ if G_n is generated by the **Disassortative Graph Algorithm** with degree sequence \mathbf{D}_n as input. An illustration of the lists S_n and T_n is displayed in Figure 2.

We will illustrate the **DGA** on the simple degree sequence $\{1, 2, 2, 3\}$, see Figure 3. Observe that in this case $z_n = 2$. Figure 3a shows the initialization state where the we have created the the lists S_n and T_n and no stubs have been connected. We start, Figure 3b, by connecting the nodes at the top of the lists, 4 and 1. Then we move down the lists, Figure 3c, and connect 4 and 2. The next step, Figure 3d, is where the specific way the algorithm ordered the stubs in both lists comes into play.

There is one stub left on the node with the largest degree, node 4. The smallest degree among the still available nodes is two. Therefore we want to connect node 4 to a node with degree two

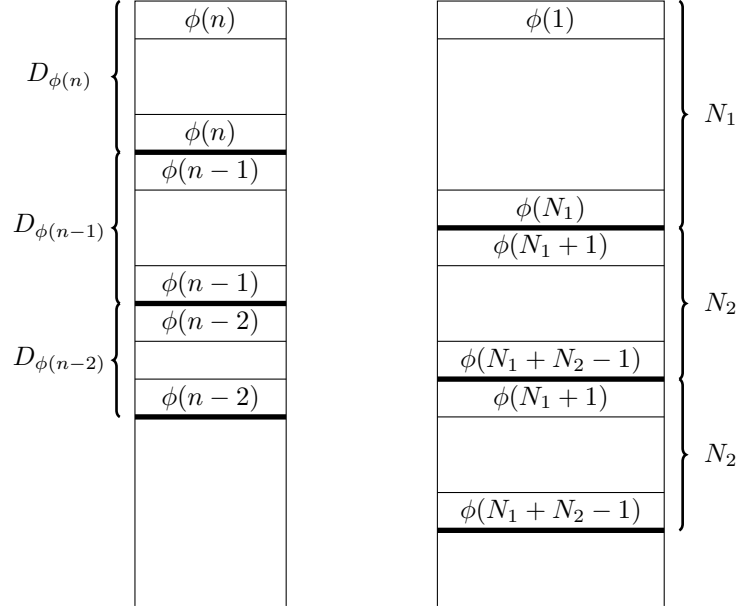


Figure 2: Top part of the two lists of stubs.

which are 2 and 3. However, since there is already an edge between 4 and 2, connecting them again will create multi-edges between these nodes. The ordering of the lists resolves this by making sure we first connected to each different node with the same degree before we can create an edge between two nodes that have already been connected. In this example we therefore connect 4 and 3.

After this step the algorithm reaches a pair of nodes that both have degree $z_n = 2$, Figure 3e. This is where we stop and pair the remaining stubs using the configuration model. Since in this specific example only nodes 2 and 3 have a stub left, we connect these, Figure 3f.

Although the DGA is defined for arbitrary degree sequences, in practice we would like to have $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ for some $\eta, \varepsilon > 0$ and distributions f and f^* . A well known algorithm for generating degree sequences with a given distribution f is by sampling the degrees i.i.d. from the distribution and then increase the last degree by 1 if the sum was not even. We will refer to this as the IID algorithm. The following lemma states that when the distribution from which the degrees are sampled has just a bit more than finite mean, the resulting degree sequence satisfies Assumption 2.1.

Lemma 3.1. *Let D be an integer valued random variable with a distribution f , such that $\mathbb{E}[D^{1+\eta}] < \infty$ for some $\eta > 0$. Denote by μ the mean of f and define $f^*(k) = \mathbb{E}[D \mathbb{1}_{\{D=k\}}] / \mu$. Then if \mathbf{D}_n is generated by the IID algorithm by sampling from f ,*

$$\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*) \quad \text{for any } \varepsilon \leq \eta / (8 + 4\eta).$$

Moreover,

$$\mathbb{P}(\Omega_n) \geq 1 - O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$.

3.2 Joint degree distribution of maximally disassortative graphs

Before we turn to analysis of the DGA it is useful to look at the empirical joint degree distribution of graphs generated by the algorithm. We will give a complete characterization of both the empirical and limiting joint degree distributions in Proposition 3.2 and Theorem 3.3, respectively.

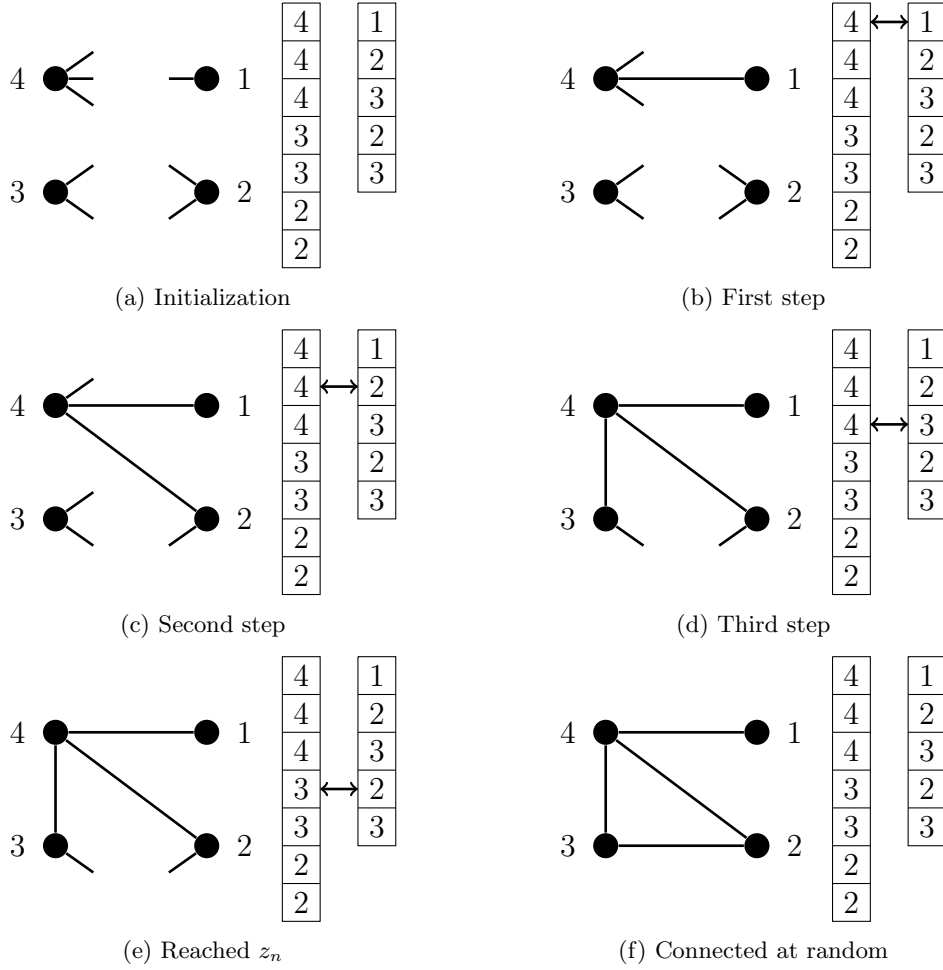


Figure 3: Example of the DGA on a simple degree sequence with $z_n = 2$.

In order to analyze the structure of the joint degree distribution we approach the algorithm from a different angle. First observe that if we are only interested in the degrees D_i and D_j for an undirected edge $i - j$, then the specific way in which the stubs are ordered by the algorithm is irrelevant for $h_n^*(k, \ell)$. This means we do not have to consider the label of the nodes to which stubs belong, only their degree. Note that the number of stubs belonging to nodes of degree k equals kN_k . Moreover, due to the symmetry in the transition to directed edges, by replacing an edge $i - j$ with edges $i \rightarrow j$ and $j \rightarrow i$, the directed structure of the graph generated by DGA can be seen as follows.

Consider the partition of the set $\{1, 2, \dots, L_n\}$ given by kN_k for $k = 0, 1, \dots$, represented as a line of length L_n partitioned into intervals of size kN_k . Now take a copy of this partitioned line, reverse it and place it below the original one, see Figure 4. Both lines can be seen as the lists of all stubs, ordered by the degree of the nodes to which they belong. For the top line the stubs are ordered, from left to right, in increasing order of the degree, while for the bottom line the degrees are in decreasing order. Then the DGA can be seen as creating directed edges $i \rightarrow j$ between the nodes i corresponding to the stubs on the bottom line and nodes j corresponding to stubs in the top line.

From this representation we observe that an edge $i \rightarrow j$ between nodes of degree $D_i = k$ and $D_j = \ell$ exists if and only if the interval corresponding to kN_k in the partitioned bottom line has an intersection with the interval corresponding to ℓN_ℓ in the partitioned upper line. In terms of

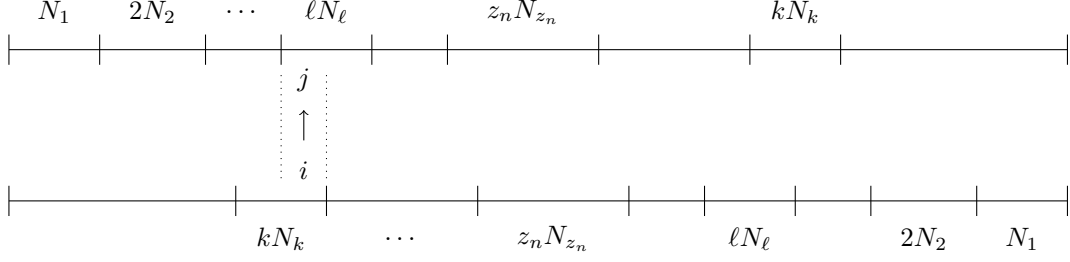


Figure 4

N_t this holds, if and only if,

$$\sum_{t=k+1}^{\infty} tN_t < \sum_{t=1}^{\ell} tN_t \quad \text{and} \quad \sum_{t=1}^{\ell-1} tN_t < \sum_{t=k}^{\infty} tN_t. \quad (15)$$

Moreover, the number of edges that connect nodes of degree k and ℓ is equal to the size of the intersection,

$$\min \left\{ \sum_{t=k}^{\infty} tN_t, \sum_{t=1}^{\ell} tN_t \right\} - \max \left\{ \sum_{t=k+1}^{\infty} tN_t, \sum_{t=1}^{\ell-1} tN_t \right\}. \quad (16)$$

This partitioned representation of both the DGA and the joint degree structure, as displayed in Figure 4, will be crucial for the analysis of the structure of maximally disassortative graphs.

First let $\psi_n(k, \ell)$ denote the indicator that there exists a directed edge $i \rightarrow j$ with $D_i = k$ and $D_j = \ell$. Then since for any $k \geq 0$,

$$\frac{1}{L_n} \sum_{t=0}^k tN_t = \frac{1}{L_n} \sum_{t=0}^k t \sum_{i=1}^n \mathbb{1}_{\{D_i=t\}} = F_n^*(k),$$

it follows from (15) that

$$\psi_n(k, \ell) := \mathbb{1}_{\{1 - F_n^*(k) < F_n^*(\ell)\}} \mathbb{1}_{\{1 - F_n^*(k-1) > F_n^*(\ell-1)\}}. \quad (17)$$

Moreover, if we let $\mathcal{E}_n(k, \ell)$ denote the average number of edges between nodes of degree k and ℓ , then (16) implies that

$$\mathcal{E}_n(k, \ell) = \min(1 - F_n^*(k-1), F_n^*(\ell)) - \max(1 - F_n^*(k), F_n^*(\ell-1)). \quad (18)$$

Summarizing we therefore have the following result.

Proposition 3.2. *Let $G_n = \text{DGA}(\mathbf{D}_n)$, for some degree sequence \mathbf{D}_n and define the functions ψ_n and \mathcal{E}_n , on the positive integers by*

$$\psi_n(k, \ell) = \mathbb{1}_{\{1 - F_n^*(k) < F_n^*(\ell)\}} \mathbb{1}_{\{1 - F_n^*(k-1) > F_n^*(\ell-1)\}} \quad \text{and} \quad (19)$$

$$\mathcal{E}_n(k, \ell) = \min(1 - F_n^*(k-1), F_n^*(\ell)) - \max(1 - F_n^*(k), F_n^*(\ell-1)). \quad (20)$$

Then,

$$h_n(k, \ell) = \psi_n(k, \ell) \mathcal{E}_n(k, \ell).$$

From Proposition 3.2 we obtain the limiting joint degree distribution as defined in (10), when $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$.

Theorem 3.3. Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ and $G_n = \text{DGA}(\mathbf{D}_n)$. In addition let $h(k, \ell)$ be as defined in (10), take $0 < \delta < \varepsilon$, $K > 0$ and define the event

$$\Xi_n = \left\{ \sum_{k, \ell=0}^{\infty} |h_n(k, \ell) - h(k, \ell)| \leq Kn^{-\delta} \right\}.$$

Then

$$\mathbb{P}(\Xi_n) \geq 1 - O\left(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)\right),$$

as $n \rightarrow \infty$.

We will use Theorem 3.3 in Section 6.5 to prove our main result, Theorem 2.1.

3.3 Properties of the Disassortative Graph Algorithm

We will now address several properties of the Disassortative Graph Algorithm. The first is concerned with the optimization problem (12).

Theorem 3.4. The Dissassortative Graph Algorithm solves (12).

This result can be explained as follows. Let \mathbf{a}_{L_n} be the list of degrees with respect to the labels of the stubs, ordered in descending order. That is

$$\mathbf{a}_{2L_n} = \left(\underbrace{D_{\phi(n)}, \dots, D_{\phi(n)}}_{D_{\phi(n)}}, D_{\phi(n-1)}, \dots, \underbrace{D_{\phi(N_1)}, \dots, D_{\phi(1)}}_{N_1} \right).$$

Then the DGA pairs the degrees a_i and a_{L_n+1-i} , which minimizes $\sum_{i \rightarrow j \in G} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j)$ and hence the DGA minimizes Spearman's rho $\tilde{\rho}(G_n)$. Observe that, in addition, the algorithm minimizes $\sum_{i \rightarrow j \in G} D_i D_j$ so that we also obtain the minimum for the s metric of the graph G , s_{\min} , as considered in [1]. Moreover, the fact that we could use an arbitrary algorithm to connect the nodes of degree z_n confirms the observation in [1] that graphs with minimal s metric are not unique with respect to their structure.

As we have already mentioned, the joint degree structure, and hence the optimality of the DGA, depends only on the degree of nodes that are connected and not on their labels. In the algorithm, however, we use an ordering for filling the lists of stubs S_n and T_n . This is to make sure that the probability that G_n^* is simple, i.e. it has no self-loops and no more than one edge between nodes i and j , converges to one as $n \rightarrow \infty$.

To understand the intuition behind the proof, consider the first time the algorithm sees a stub belonging to a node i in the list S_n with degree $D_i > z_n$. Then node i will be connected to the nodes corresponding to the next D_i stubs in T_n . Now consider such a stub, belonging to node j . Then there will be more than one edge $i - j$ if and only if there is more than one stub belonging to node j among the D_i stubs in T_n , which can only happen when $D_i > N_{D_j}$. Since the degree of nodes in T_n is bounded by z_n , we have that N_{D_j} scales as n , while the maximal degree is $o(n)$, since f has finite mean. Therefore, the event $D_i > N_{D_j}$ for $D_i > z_n$ and $D_j \leq z_n$ has vanishing probability. We hence have the following result, the details of the proof can be found in Section 6.3.

Proposition 3.5. Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$, $G_n = \text{DGA}(\mathbf{D}_n)$ and denote by \mathcal{S}_n^* the event that G_n^* is simple, then

$$\mathbb{P}(\mathcal{S}_n^*) \geq 1 - O\left(n^{-\varepsilon} + n^{-1/2} + n^{-\eta/2} + \mathbb{P}(\Omega_n^c)\right),$$

as $n \rightarrow \infty$.

This proposition implies that the simplicity of the graph G_n , generated by the **Disassortative Graph Algorithm**, solely depends on the simplicity of G_n^z , constructed in Step 19. Now consider the degree sequence \mathbf{D}_n^z corresponding to the remaining stubs, obtained in Step 18 and observe that these degrees are uniformly bounded by z_n . Take $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ and let z be the median of F^* , i.e. the unique integer such that

$$F^*(z) \geq \frac{1}{2} \quad \text{and} \quad F^*(z-1) < \frac{1}{2}. \quad (21)$$

We can show that

$$\lim_{n \rightarrow \infty} \mathbb{P}(z_n \leq z+1) = 1,$$

see the proof of Proposition 3.5 in Section 6.3. Therefore, if we define the event $A_n = \{z_n \leq z+1\}$, then conditioned on A_n the degrees in \mathbf{D}_n^z are bounded by $z+1$. Hence, if we connect these stubs using the configuration model, and let \mathcal{S}_n^z denote the event that G_n^z is simple, then it follows, see e.g. [15] Theorem 7.12, that there exist a constant $\mathfrak{s} > 0$, such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{S}_n) = \lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{S}_n^z, A_n) + \mathbb{P}(A_n^c) = \mathfrak{s}.$$

From this and Proposition 3.5 we obtain the following corollary.

Corollary 3.6. *Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}$, $G_n = \text{DGA}(\mathbf{D}_n)$ and denote by \mathcal{S}_n the event that G_n is simple. Then there exists a constant $\mathfrak{s} > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{S}_n) = \mathfrak{s}.$$

Note that by Lemma 3.1 it follows that if D is an integer valued random variable that satisfies, for some $\eta > 0$,

$$\nu := \mathbb{E}[D] < \infty \quad \text{and} \quad \mathbb{E}[D^{1+\eta}] < \infty,$$

then a degree sequence \mathbf{D}_n generated by the IID algorithm satisfies $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}$, for any $0 < \varepsilon < \eta/(8+4\eta)$, while

$$\mathbb{P}(\Omega_n) \geq 1 - O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$. Therefore, if we want to generate maximally disassortative graphs with limit degree density f , we can first generate a degree sequence using the IID algorithm, by sampling from f , and then connect the nodes using the DGA. From Corollary 3.6 it now follows that, in order to generate maximally disassortative simple graphs, we could repeat steps 13 to 19 in the **Disassortative Graph Algorithm** until the resulting graph is simple.

4 Spearman's rho and the tail of the degree distribution

We will now investigate the influence of the degree distribution on the value of Spearman's rho, on maximally disassortative graphs, i.e. graphs generated by the DGA. We will show that the tail of the distribution does not influence this value. This is achieved by transforming a given degree distribution, such that the asymptotic behavior of the tail of this distribution is preserved, while we increase the probability mass of the corresponding size-biased degree distribution at one.

Let us start by considering a degree distributions f , for which the size-biased distributions f^* satisfies $f^*(1) \geq 1/2$, i.e. f^* has median 1. Observe that in this case we have

$$h(k, \ell) = \begin{cases} f^*(k) & \text{if } k > 1 \text{ and } \ell = 1 \\ 2f^*(1) - 1 & \text{if } k = 1 \text{ and } \ell = 1 \\ f^*(\ell) & \text{if } k = 1 \text{ and } \ell > 1 \\ 0 & \text{else.} \end{cases}$$

Hence, if D_*, D^* have joint distribution h , as defined in (10), then

$$\begin{aligned}
\mathbb{E}[\mathcal{F}^*(D_*)\mathcal{F}^*(D^*)] &= \sum_{k,\ell=1}^{\infty} \mathcal{F}^*(k)\mathcal{F}^*(\ell)h(k,\ell) \\
&= f^*(1)^2(2f^*(1) - 1) + 2f^*(1) \sum_{\ell=2}^{\infty} \mathcal{F}^*(\ell)f^*(\ell) \\
&\geq f^*(1)^2(2f^*(1) - 1) + 4f^*(1)^2 \sum_{\ell=2}^{\infty} f^*(\ell) \\
&= f^*(1)^2(2f^*(1) - 1) + 4f^*(1)^2(1 - f^*(1)) \\
&= 3f^*(1)^2 - 2f^*(1)^3,
\end{aligned}$$

where we used, see (4), that $\mathcal{F}^*(\ell) \geq 2f^*(1)$ for all $\ell > 1$. From this it follows that whenever $f^*(1) \geq 1/2$ and D_*, D^* have joint distribution h ,

$$3\mathbb{E}[\mathcal{F}^*(D_*)\mathcal{F}^*(D^*)] - 3 \geq 9f^*(1)^2 - 6f^*(1)^3 - 3. \quad (22)$$

Since the function on the right side of (22) is strictly monotonically increasing and is 0 when $f^*(1) = 1$, it follows that the limit of Spearman's rho on maximally disassortative graphs can be bounded from below by a value that is arbitrary close to 0, if $f^*(1)$ is large enough. Moreover, using that the h is the joint degree distribution of graphs with minimal value of $\tilde{\rho}$, we have the following result.

Proposition 4.1. *Let f and f^* be such that $f^*(1) \geq \frac{1}{2}$ and $G_n \in \mathcal{G}_{\eta,\varepsilon}(f, f^*)$. Then, for any $0 < \delta < \min(\varepsilon, 1/2)$ and $K > 0$,*

$$\mathbb{P}(\tilde{\rho}(G_n) \geq 9f^*(1)^2 - 6f^*(1)^3 - 3 - Kn^{-\delta}) \geq 1 - O(n^{-\varepsilon+\kappa} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$, where $\kappa = (\varepsilon + \delta)/2$.

Given $0 < \delta < 1$, we will now describe a construction that transforms any given distribution f with support on the positive integers, into a distribution f_δ , with support on the positive integers, such that $f_\delta(1) = \delta$ and

$$\lim_{k \rightarrow \infty} \frac{1 - F_\delta(k)}{1 - F(k)} = 1, \quad (23)$$

where F and F_δ are the cumulative distribution functions of f and f_δ , respectively.

To see that f_δ defines a probability density function we compute

$$\begin{aligned}
\sum_{t=1}^{\infty} f_\delta(t) &= \delta + \sum_{t=2}^{K_\delta+1} f_\delta(t) + \sum_{t=K_\delta+2}^{\infty} f(t) \\
&= \delta + x + f(K_\delta + 1) + 1 - F(K_\delta + 1) = 1.
\end{aligned}$$

Moreover, since $f_\delta(k) = f(k)$ for all $k > K_\delta$ it follows that F_δ satisfies (23). We will refer to f_δ as the δ -transform of f .

With this transformation we can now transform a given distribution f , to get a distribution f_ρ whose size-biased distribution f_ρ^* satisfies

$$9f_\rho^*(1)^2 - 6f_\rho^*(1)^3 - 3 > \rho,$$

without affecting the asymptotic behavior of the tail of the original distribution f . It then follows from Proposition 4.1 that for any sequence of graphs $G_n \in \mathcal{G}_{\eta,\varepsilon}(f_\rho, f_\rho^*)$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tilde{\rho}(G_n) > \rho) = 1,$$

Algorithm 2 δ -transform of a probability density f

```
1: Given: a probability density  $f$ , corresponding cdf  $F$  and  $0 < \delta < 1$ 
2: Let  $K_\delta$  be the smallest integer such that  $F(K_\delta) > \delta$ .
3: Set  $x = F(K_\delta) - \delta > 0$  and  $f_\delta(1) = \delta$ .
4: if  $K_\delta = 1$  then
5:   Set  $f_\delta(2) = f(2) + x$ 
6: else
7:   for  $2 \leq k \leq K_\delta$  do
8:     Set  $f_\delta(k) = x/(K_\delta - 1)$ 
9:   end for
10:  Set  $f_\delta(K_\delta + 1) = f(K_\delta + 1)$ 
11: end if
12: for  $k > K_\delta$  do
13:   Set  $f_\delta(k) = f(k)$ 
14: end for
15: Output: Probability density  $f_\delta$ 
```

which proves Theorem 2.2. The details can be found in Section 6.5.

The construction we use for creating the adversary degree distribution f_ρ has one downside. In order to construct degree distributions such that $\tilde{\rho}(G_n)$ is arbitrary close to zero, the value of $f^*(1)$ should be arbitrary close to 1. Therefore, these distributions might not resemble real-world situations. The reason for this downside is that the construction is based on the very crude lower bound (22) on Spearman's rho, for which we had to assume $f^*(1) \geq 1/2$.

As we mentioned in Section 2.3, Theorem 2.2 states that the minimal value of Spearman's rho and not determined by the tail of the distribution.

Now let F be regularly varying with exponent $\gamma > 1$ and slowly varying function \mathcal{L} , see (13). Pick any $-1 < \rho < 0$ and let F_ρ be the transformed distribution, given by Theorem 2.2. We will show that F_ρ is again regularly varying with exponent γ . Note that for this it is enough to show that $(1 - F_\rho(x))x^\gamma$ is slowly varying. To this end fix $t > 0$ and write

$$\begin{aligned} \frac{1 - F_\rho(tk)}{t^{-\gamma}(1 - F_\rho(k))} &= \left(\frac{1 - F_\rho(tk)}{1 - F(tk)} \right) \left(\frac{1 - F(k)}{1 - F_\rho(k)} \right) \left(\frac{1 - F(tk)}{t^{-\gamma}(1 - F(k))} \right) \\ &= \left(\frac{1 - F_\rho(tk)}{1 - F(tk)} \right) \left(\frac{1 - F(k)}{1 - F_\rho(k)} \right) \frac{\mathcal{L}(tk)}{\mathcal{L}(k)}. \end{aligned}$$

The product of the first two terms converge to 1, as $k \rightarrow \infty$, by Theorem 2.2, while this holds for the last term since \mathcal{L} is slowly varying. Summarizing, we have

$$\lim_{k \rightarrow \infty} \frac{1 - F_\rho(tk)}{t^{-\gamma}(1 - F_\rho(k))} = 1,$$

which proves that $(1 - F_\rho(x))x^\gamma$ is slowly varying and hence F_ρ is regularly varying with exponent γ . This proves Corollary 2.3.

5 Spearman's rho on maximal disassortative graphs.

We will now use numerical experiments to illustrate the behavior of Spearman's rho for two types of degree distributions, regularly varying and Poisson. Each of these types has a parameter that can serve as a proxy for the way in which the mass of the probability density functions is distributed over their support. For the regularly varying distributions this is the exponent γ , while for the Poisson distribution it is the mean λ . We will refer to these as the parameters of the distribution.

For the simulations we generated degree sequences \mathbf{D}_n by sampling from the given distribution, using the IID algorithm, for different sizes n and values for the parameters. We then generated

graphs G_n using the DGA. For each combination of size and parameter, we generated 10^3 graphs in this manner and computed $\rho(G_n)$, as defined in (6), on each of them. This gives us 10^3 samples of Spearman's rho on maximal disassortative graphs with the given size and degree distribution.

To analyze the speed of convergence of $\rho(G_n)$ we computed for each combination of size and parameter

$$X_n := |\rho(G_n) - \mathbb{E}'[\rho(G_n)]|,$$

where \mathbb{E}' denotes the empirical mean, based on the 10^3 realizations per such combination. We then plotted the empirical inverse cumulative distribution of X_n for different sizes $n = 10^4, 10^5, 10^6$ and 10^7 . The results are shown in Figure 5 and Figure 6.

In addition, to investigate the limit of Spearman's rho in maximally disassortative graphs, we computed $\mathbb{E}'[\rho(G_n)]$, with $n = 10^7$, for several values of the parameter of the distribution. We then plotted these values with respect to the parameter in Figure 7.

We will now describe the specific distributions we used for the simulations and discuss the results.

5.1 Scale-free degree distribution

Let X have a Pareto distribution with scale 1 and shape $\gamma > 1$, i.e.

$$f_X(t) = \begin{cases} \gamma t^{-1-\gamma} & \text{if } t \geq 1 \\ 0 & \text{else,} \end{cases} \quad 1 - F_X(t) = \begin{cases} t^{-\gamma} & \text{if } t \geq 1 \\ 1 & \text{else,} \end{cases}$$

and define $D = \lfloor X \rfloor$. Then we have that $1 - F(k) = 1 - F_X(k+1)$, so that F is regularly varying with exponent $\gamma > 1$, while

$$f(k) = F(k) - F(k-1) = k^{-\gamma} - (k+1)^{-\gamma}. \quad (24)$$

Standard calculations yield that $\sum_{k=0}^{\infty} k f(k) = \zeta(\gamma)$, where ζ is the Riemann zeta function. Therefore we have that

$$f^*(k) = \frac{k f(k)}{\zeta(\gamma)},$$

so that $f^*(1) = (1 - 2^{-\gamma})/\zeta(\gamma)$ which is increasing in γ . Moreover $9f^*(1)^2 - 6f^*(1)^3 - 3 > -1$ for all $\gamma \geq 2.5$, which places it in the class of adversary distributions we considered in the previous section.

From Figure 5 we see that $\rho(G_n)$ is already strongly concentrated around its mean when $n = 10^5$. Even when the degree distribution has infinite variance ($\gamma = 1.5$) we have that $X_n \leq 0.025$, with high probability, for graphs of size $n = 10^5$. This shows, complementary to Theorem 2.1, that the DGA performs very well in practice with respect to the convergence of Spearman's rho to the minimal achievable value $\rho(D_*, D^*)$.

Interestingly, the simulations suggest that the concentration of $\rho(G_n)$ around its mean for graphs of small size becomes tighter when γ decreases. Compare, for instance, the plots for $n = 10^4$ in the Figure 5a - 5c.

In Figure 7a we plotted the empirical average of $\rho(G_n)$ against the parameter γ of the degree density (24). Observe that in contrary to the lower bound related to $f^*(1)$, we clearly see that Spearman's rho is strongly increasing as a function of γ and $\rho(G_n) > -0.8$ for $\gamma > 2$. Therefore it follows by Theorem 2.1 that the rank-correlation measure Spearman's rho on any graph with degree distribution (24) and $\gamma > 2$ will not have a value smaller than -0.8 . Moreover, when $\gamma \geq 2.5$ we see that $\mathbb{E}'[\rho(G_n)] > -0.5$. Since this is a lower bound for Spearman's rho on any graph with degree density (24), a consequence could be that even if such graphs have a very disassortative joint degree structure they could potentially be classified differently.

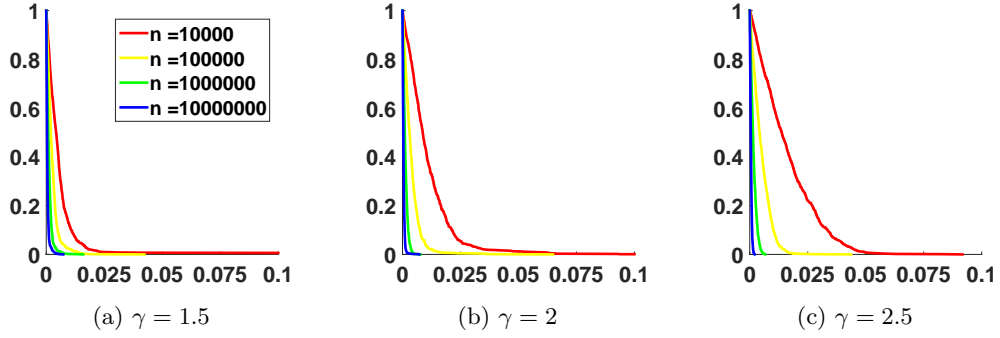


Figure 5: Plot of the inverse cdf of X_n for graphs of different sizes and degree distribution (24), generated by the DGA, for three different choices of γ .

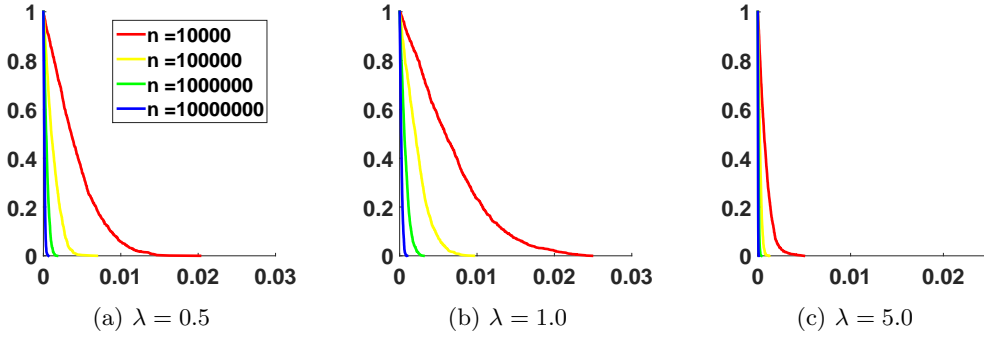


Figure 6: Plot of the inverse cdf of X_n for graphs of different sizes and Poisson degree distribution, generated by the DGA, for three different choices of λ .

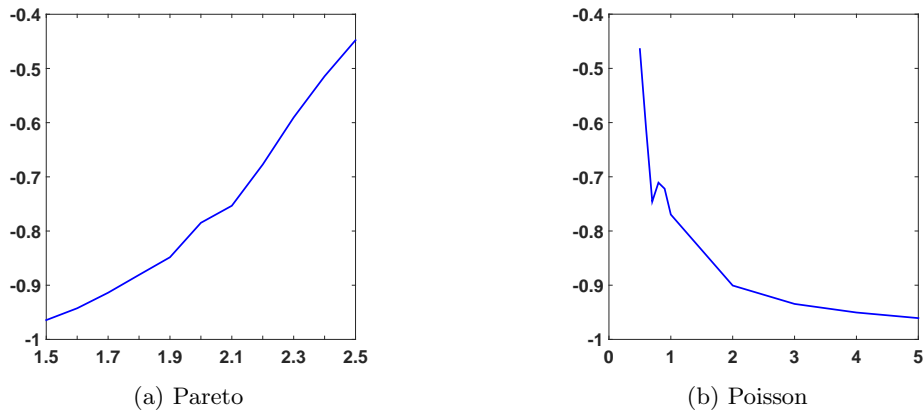


Figure 7: Plot of the empirical average of $\rho(G_n)$ for graph of size 10^7 and degree distributions (24) and Poisson, generated by the DGA, for different values of, respectively, γ and λ .

5.2 Poisson degree distribution

Let X be a Poisson random variable with mean λ and denote its probability by f . Then it follows that $f^*(k) = f(k-1)$. Hence $f^*(1) = e^{-\lambda}$ is a decreasing function of λ and $9f^*(1)^2 - 6f^*(1)^3 - 3 > -1$ for at least all $\lambda \leq 0.4$. This is opposite to the degree distribution (24), where $f^*(1)$ was an increasing function of the parameter γ . This is reflected in Figure 7b, where we see that $\mathbb{E}'[\rho(G_n)]$ decreases with λ . Here we again see that the shape of the degree distribution strongly influences the value of $\rho(G_n)$ for maximally disassortative graphs, and hence the minimal value that Spearman's rho can attain for any graph with this degree distribution. Note that, in contrast to the case with the regularly varying distribution, $\rho(G_n)$ is not monotonic with respect to λ . This could be due to the fact that the Poisson density is non-monotonic, while the density (24) is monotonically decreasing.

In addition we also observe that, similar to the previous setting, the DGA performs very well with respect to the convergence of $\rho(G_n)$. Already for very reasonable sizes, $n \geq 10^5$, the deviations around the mean are, with high probability, smaller than 0.02 for all three values of λ .

5.3 Important observations and insights

The main observation from the simulations that we did is that the distribution of the mass of the degree probability density is of crucial importance for the minimal value that Spearman's rho can attain. Moreover, it seems that already for very reasonable degree distributions this minimum is much larger than -1 . Therefore, one should be careful when classifying a network as not being very disassortative when a small negative value of Spearman's rho is computed.

The simulations suggest something even stronger. For this, consider the probability density (24) and observe that if we increase γ then the atoms at the end of this density lose mass, while those at the beginning gain mass. In this way we can use the parameter γ to 'shift' mass between the head and tail of the distribution. The mean, λ , of the Poisson can be used in a similar way, although in this case we need to decrease λ in order to move mass towards the head. For both distributions we see that, as the mass of the probability density f is moved towards the tail (decreasing γ /increasing λ), the value of $\rho(G_n)$ in maximally disassortative graphs with this degree distribution decreases and seems to approach -1 . On the other hand, as we move more mass to the head of the probability density f (increasing γ /decreasing λ) the minimal value of Spearman's rho increases and seems to go to zero.

6 Proofs

Here we prove the results stated in this paper.

6.1 Generating degree sequences $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$

Proof of Lemma 3.1. We remark that altering the last degree by at most 1, to make the sum even, constitutes a correction term of order n^{-1} . Hence we will consider the degrees D_i as i.i.d. samples from D .

Now fix $\varepsilon \leq \eta/(8 + 4\eta)$ and define the events

$$A_n = \left\{ \sum_{k=0}^{\infty} \left| \sum_{t=0}^k f_n(t) - f(t) \right| \leq n^{-\eta/(2+2\eta)} \right\}$$

$$B_n = \left\{ \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \leq n^{-\varepsilon} \right\}$$

and notice that $\mathbb{P}(\Omega_n^c) \leq \mathbb{P}(A_n^c) + \mathbb{P}(A_n \cap B_n^c)$. For the first term we have, using Markov's inequality,

$$\mathbb{P}(A_n^c) \leq n^{\frac{\eta}{2+2\eta}} \mathbb{E}[d_1(f_n, f)] \leq O\left(n^{-\frac{\eta}{2+2\eta}}\right) = O\left(n^{-\varepsilon}\right),$$

as $n \rightarrow \infty$, where the second inequality follows from [4, Proposition 4.2] and the last since $-\eta/(2+2\eta) < -\eta/(8+4\eta)$. Hence, we need to show that, as $n \rightarrow \infty$

$$\mathbb{P}(A_n \cap B_n^c) \leq O(n^{-\varepsilon}).$$

For this we compute that,

$$\begin{aligned} \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| &= \sum_{k=0}^{\infty} \left| \frac{1}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i=k\}} - \frac{\mathbb{E}[D \mathbb{1}_{\{D=k\}}]}{\mu} \right| \\ &\leq \sum_{k=0}^{\infty} \left| \frac{1}{L_n} - \frac{1}{\mu n} \right| \sum_{i=1}^n D_i \mathbb{1}_{\{D_i=k\}} \\ &\quad + \sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n D_i \mathbb{1}_{\{D_i=k\}} - \mathbb{E}[D \mathbb{1}_{\{D=k\}}] \right| \\ &= \frac{|L_n - \mu n|}{\mu n} + \sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n X_{ik} \right|, \end{aligned}$$

where we defined $X_{ik} = D_i \mathbb{1}_{\{D_i=k\}} - \mathbb{E}[D \mathbb{1}_{\{D=k\}}]$. Now observe that conditioned on A_n we have that

$$\begin{aligned} \frac{|L_n - \mu n|}{\mu n^{1-\varepsilon}} &\leq \frac{n^{1-\eta/(2+2\eta)}}{\mu n^{1-\varepsilon}} \leq \mu^{-1} n^{\varepsilon-\eta/2(1+\eta)} \\ &\leq \mu^{-1} n^{-\eta/4(1+\eta)} \leq \mu^{-1} n^{-\eta/4(1+2\eta)}, \end{aligned}$$

so that

$$\mathbb{P}(A_n \cap B_n^c) \leq O(n^{-\varepsilon}) + \mathbb{P}\left(\sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n X_{ik} \right| > n^{-\varepsilon}\right),$$

as $n \rightarrow \infty$.

To analyze the last probability take $a_n = \lfloor n^{2\varepsilon/\eta} \rfloor$. Then,

$$\begin{aligned} \mathbb{P}\left(\sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n X_{ik} \right| > n^{-\varepsilon}\right) &\leq \frac{1}{\mu n^{1-\varepsilon}} \sum_{k=0}^{a_n} \mathbb{E} \left[\left| \sum_{i=1}^n X_{ik} \right|^2 \right]^{1/2} \\ &\quad + \frac{1}{\mu n^{1-\varepsilon}} \sum_{k=a_n+1}^{\infty} \mathbb{E} \left[\left| \sum_{i=1}^n X_{ik} \right| \right] \\ &\leq \frac{1}{\mu n^{1/2-\varepsilon}} \sum_{k=0}^{a_n} \text{Var}(X_{1k})^{1/2} + \frac{n^{\varepsilon}}{\mu} \sum_{k=a_n+1}^{\infty} \mathbb{E}[|X_{1k}|] \\ &\leq \frac{1}{\mu n^{1/2-\varepsilon}} \sum_{k=0}^{a_n} k + \frac{2n^{\varepsilon}}{\mu} \sum_{k=a_n+1}^{\infty} \mathbb{E}[D \mathbb{1}_{\{D=k\}}] \\ &\leq \frac{a_n(a_n+1)}{2\mu n^{1/2-\varepsilon}} + \frac{2n^{\varepsilon}}{\mu} \mathbb{E}[D \mathbb{1}_{\{D>a_n\}}] \\ &\leq \frac{a_n(a_n+1)}{2\mu n^{1/2-\varepsilon}} + 2n^{\varepsilon} a_n^{-\eta} \\ &= O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$. Here, for the last line, we used that

$$\frac{4\varepsilon}{\eta} - \frac{1}{2} + \varepsilon \leq \frac{1}{2+\eta} - \frac{1}{2} + \frac{\eta}{8+4\eta} = -\frac{\eta}{8+4\eta} \leq -\varepsilon,$$

so that

$$n^{\varepsilon-\frac{1}{2}} a_n^2 = O\left(n^{\frac{4\varepsilon}{\eta}-\frac{1}{2}+\varepsilon}\right) = O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$. □

6.2 Optimality of DGA

Theorem 3.4 is a consequence of the following lemma.

Lemma 6.1. *Consider a sequence $0 \leq a_1 \leq \dots \leq a_m$ and let \mathcal{P}_m denote the set of permutations of $\{1, \dots, m\}$. Then*

$$\min_{\sigma \in \mathcal{P}_m} \sum_k a_k a_{\sigma(k)} = \sum_k a_k a_{m-k+1}$$

and this minimum is achievable for a permutation σ if and only if

$$a_{\sigma(1)} \geq a_{\sigma(2)} \geq \dots \geq a_{\sigma(n)}.$$

Proof.

[\Rightarrow] If $a_{\sigma(1)} \geq \dots \geq a_{\sigma(n)}$ then $\sum_k a_k a_{\sigma(k)} = \sum_k a_k a_{m-k+1}$

[\Leftarrow] Assume that $\sigma = \arg \min_{\sigma \in S_m} \sum_k a_k a_{\sigma(k)}$ but there exist $a_i < a_j$ such that $a_{\sigma(i)} < a_{\sigma(j)}$. Consider $\sigma^* = \sigma \cdot (ij)$ then $\sum_k a_k a_{\sigma(k)} - \sum_k a_k a_{\sigma^*(k)} = (a_i - a_j)(a_{\sigma(i)} - a_{\sigma(j)}) > 0$ which contradicts the initial assumption. \square

Proof of Theorem 3.4. Consider a degree sequence \mathbf{D}_n , rank it in ascending order and let $\phi(k)$ denotes the node with rank k among this degree sequence, as defined in the description of the DGA. Now define the sequence \mathbf{a}_{L_n} by

$$a_k = \mathcal{F}_n^*(D_{\phi(i)}) \quad \text{for all } \sum_{t=1}^{i-1} D_{\phi(t)} < k \leq \sum_{t=1}^i D_{\phi(t)}, \quad (25)$$

where we use the convention that $\sum_{t=1}^0 D_{\phi(t)} = 0$. With this definition, the sequence \mathbf{a}_{L_n} looks as follows

$$a_1 \leq \dots \leq a_k \leq \dots = \underbrace{\mathcal{F}_n^*(D_{\phi(1)}) \leq \dots \leq \mathcal{F}_n^*(D_{\phi(1)})}_{D_{\phi(1)}} \leq \underbrace{\mathcal{F}_n^*(D_{\phi(2)}) \leq \dots \leq \mathcal{F}_n^*(D_{\phi(2)})}_{D_{\phi(2)}} \leq \dots$$

Next, we note that for each graph $G \in \mathcal{G}(\mathbf{D}_n)$ there exists a permutation σ_G of such that

$$\sum_{i \rightarrow j \in G} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j) = \sum_{k=1}^{L_n} a_k a_{\sigma_G(k)}.$$

Any directed graph, has a corresponding permutation σ of $\{1, \dots, L_n\}$ which defines how the outbound and inbound stubs of the bi-degree sequence are paired to obtain the graph. However, not every such permutation corresponds to a graph which is the bi-directed version of an undirected graph, i.e. for each edge $i \rightarrow j$ there is exactly one edge $j \rightarrow i$. Therefore let $\mathcal{P}(\mathbf{D}_n)$ denote the set of all permutations of $\{1, \dots, L_n\}$ which do corresponds to an undirected graph, in its directed representation. Then the optimization problem (12) is equivalent to the following problem

$$\min_{\sigma \in \mathcal{P}(\mathbf{D}_n)} \sum_{k=1}^{L_n} a_k a_{\sigma^*(k)}. \quad (26)$$

Now, recall the partitioned representation of the DGA we introduced in Section 3.2, see Figure 4. From this description of the algorithm it is not hard to see that, if \mathbf{a}_{L_n} is defined by (25), then there exists a permutation σ^* with the property that

$$a_{\sigma^*(1)} \geq a_{\sigma^*(2)} \geq \dots \geq a_{\sigma^*(L_n)},$$

such that the DGA pairs the stubs corresponding to a_i and $a_{\sigma^*(i)}$. Therefore, Lemma 6.1 implies that

$$\sum_{k=1}^{L_n} a_k a_{\sigma^*(k)} = \min_{\sigma \in \mathcal{P}_{L_n}} \sum_{k=1}^{L_n} a_k a_{\sigma(k)},$$

where \mathcal{P}_{L_n} denotes the set of all permutations of $\{1, \dots, L_n\}$. Since $\mathcal{P}(\mathbf{D}_n) \subseteq \mathcal{P}_{L_n}$, this implies that

$$\sum_{k=1}^{L_n} a_k a_{\sigma^*(k)} = \min_{\sigma \in \mathcal{P}(\mathbf{D}_n)} \sum_{k=1}^{L_n} a_k a_{\sigma(k)},$$

which proves that the DGA solves (26) and hence it solves (12) \square

6.3 Simplicity of G_n^*

Proof of Proposition 3.5. Let z_n and z be defined as in (14) and (21), respectively, and define the event

$$A_n = \{z_n \leq z + 1\}.$$

Then, by definition of z_n , we have that $F^*(z + 1) > 1/2$ and hence

$$\begin{aligned} \mathbb{P}(z_n > z + 1, \Omega_n) &\leq \mathbb{P}\left(F_n^*(z + 1) < \frac{1}{2}, \Omega_n\right) \\ &\leq \mathbb{P}\left(F^*(z + 1) - \frac{1}{2} < |F^*(z + 1) - F_n^*(z + 1)|, \Omega_n\right) \\ &\leq \frac{\mathbb{E}[|F^*(z + 1) - F_n^*(z + 1)| \mathbb{1}_{\{\Omega_n\}}]}{F^*(z + 1) - \frac{1}{2}} \\ &\leq O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$. Therefore, if we define $\Lambda_n = A_n \cap \Omega_n$, it is enough to show that

$$1 - \mathbb{P}(\mathcal{S}_n, \Lambda_n) \leq O(n^{-\varepsilon} + n^{-1/2} + n^{-\eta/2}).$$

In order to analyze this probability, note that by construction there are no self-loops in G_n^* . Moreover, a node i with $D_i < z_n$ can only have more than one edge to a node j when $D_j > N_{D_i}$. Hence, when G_n^* is not simple it means that for some $1 \leq k \leq z_n$ we must have that $0 < N_k < \max_{1 \leq j \leq n} D_j$, hence

$$(\mathcal{S}_n^c)^c \subseteq \bigcup_{k=1}^{z_n} \left\{ 0 < N_k < \max_{1 \leq j \leq n} D_j \right\}.$$

Therefore, if we denote $f_{\min} = \min_{1 \leq k \leq z} f(k) > 0$, it follows from the union bound that

$$\begin{aligned} 1 - \mathbb{P}(\mathcal{S}_n, \Lambda_n) &\leq \sum_{k=1}^{z+1} \mathbb{P}\left(0 < N_k < \max_{1 \leq j \leq n} D_j, \Lambda_n\right) \\ &= \sum_{k=1}^{z+1} \mathbb{P}\left(0 < f_n(k) < \frac{\max_{1 \leq j \leq n} D_j}{n}, \Lambda_n\right) \\ &\leq \sum_{k=1}^{z+1} \mathbb{P}\left(f(k) - n^{-\varepsilon} < \frac{\max_{1 \leq j \leq n} D_j}{n}, \Lambda_n\right) \\ &\leq \sum_{k=1}^{z+1} \mathbb{P}\left(f_{\min} < \frac{\max_{1 \leq j \leq n} D_j}{n} + n^{-\varepsilon}, \Lambda_n\right) \end{aligned}$$

$$\begin{aligned}
&\leq \frac{(z+1)n^{-1}\mathbb{E}[\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{\Lambda_n\}}] + (z+1)n^{-\varepsilon}}{f_{\min}} \\
&\leq \frac{(z+1)\mathbb{E}[\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{D_j > \sqrt{n}\}} \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\
&\quad + \frac{(z+1)n^{-1/2}}{f_{\min}} + \frac{(z+1)n^{-\varepsilon}}{f_{\min}}.
\end{aligned}$$

The last probability is $O(n^{-1/2} + n^{-\varepsilon})$, as $n \rightarrow \infty$. We will now show that the other probability is $O(n^{-\varepsilon} + n^{-\eta/2})$. For this we note that

$$\frac{\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{D_j > \sqrt{n}\}}}{n} \leq \frac{1}{n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i > \sqrt{n}\}} = \frac{L_n}{n} (1 - F_n^*(\sqrt{n})),$$

and

$$1 - F^*(\sqrt{n}) = \mathbb{E}[D \mathbb{1}_{\{D > \sqrt{n}\}}] \leq n^{-\eta/2} \mathbb{E}[D^{1+\eta}].$$

Therefore we obtain that, as $n \rightarrow \infty$,

$$\begin{aligned}
\frac{\mathbb{E}[\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} &\leq \frac{\mathbb{E}[L_n |F_n^*(\sqrt{n}) - F^*(\sqrt{n})| \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\
&\quad + \frac{\mathbb{E}[2L_n(1 - F^*(\sqrt{n})) \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\
&\leq \frac{\mathbb{E}[(\nu n + n^{1-\varepsilon}) \sup_{k \geq 0} |F_n^*(k) - F^*(k)| \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\
&\quad + \frac{\mathbb{E}[2(\nu n + n^{1-\varepsilon})(1 - F^*(\sqrt{n})) \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\
&\leq \frac{(\nu + n^{-\varepsilon})n^{-\varepsilon}}{nf_{\min}} + \frac{2(\nu + n^{-\varepsilon})n^{-\eta/2} \mathbb{E}[D^{1+\eta}]}{f_{\min}} \\
&\leq O(n^{-\varepsilon} + n^{-\eta/2}),
\end{aligned}$$

which completes the proof. \square

6.4 Joint degree distribution

Here we will address the convergence of the empirical joint degree density $h_n(k, \ell)$, as defined in Proposition 3.2, to the density $h(k, \ell)$ as defined in (10).

We will use two technical lemmas, which deal with the difference between the functions ψ_n and ψ , and \mathcal{E}_n and \mathcal{E} .

Lemma 6.2. *Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$. Then, for any $k, \ell \geq 0$, $0 < \delta < \varepsilon$ and $K > 0$*

$$\mathbb{P}(|\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > Kn^{-\delta}, \Omega_n) \leq O(n^{-\varepsilon+\delta}).$$

Lemma 6.3. *Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$. Then, for any $k, \ell \geq 0$, $K > 0$ and $0 < \delta < \varepsilon$,*

$$\mathbb{P}(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > Cn^{-\delta}, \Omega_n) \leq O(n^{-\varepsilon+\delta}).$$

The proof of both lemmas is postponed till the end of this section. We will first give the proof of Theorem 3.3.

Proof of Theorem 3.3. Let p be the smallest integer satisfying

$$1 - F^*(p) < f^*(1), \quad (27)$$

and define p_n as the smallest integer that satisfies

$$1 - F_n^*(p_n) < f_n^*(1).$$

Then we have that $\mathbb{P}(p_n = p)$ converges to one, since

$$\begin{aligned} \mathbb{P}(p \neq p_n, \Omega_n) &= \mathbb{P}(1 - F_n^*(p) \geq f_n^*(1), \Omega_n) \\ &\leq \mathbb{P}((F^*(p) - F_n^*(p)) + (f^*(1) - f_n^*(1)) > f^*(1) - 1 + F^*(p), \Omega_n) \\ &\leq \mathbb{P}(|F^*(p) - F_n^*(p)| > f^*(1) - 1 + F^*(p), \Omega_n) \\ &\quad + \mathbb{P}(|f^*(1) - f_n^*(1)| > f^*(1) - 1 + F^*(p), \Omega_n) \\ &\leq \frac{\mathbb{E}[|F^*(p) - F_n^*(p)| \mathbb{1}_{\{\Omega_n\}}]}{f^*(1) - 1 + F^*(p)} + \frac{\mathbb{E}[|f^*(1) - f_n^*(1)| \mathbb{1}_{\{\Omega_n\}}]}{f^*(1) - 1 + F^*(p)} \\ &\leq \frac{2n^{-\varepsilon}}{f^*(1) - 1 + F^*(p)} \leq O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$, where we used that by definition of p it holds that $f^*(1) - 1 + F^*(p) > 0$. Therefore, if we define the event $P_n = \{p = p_n\}$ and let $\Lambda_n = P_n \cap \Omega_n$, then

$$\mathbb{P}(\Lambda_n) \geq 1 - O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)),$$

so that for Theorem 3.3 it is enough to show that

$$\mathbb{P}(\Xi_n^c, \Lambda_n) \leq O(n^{-\varepsilon+\delta}), \quad (28)$$

as $n \rightarrow \infty$.

Now, observe that p_n is the smallest degree such that nodes i with degree $D_i > p_n$ will be connected to nodes with degree 1, by the DGA, while p is the corresponding degree for the limit distribution. Therefore we have

$$\psi(k, \ell) = \begin{cases} 1 & \text{for all } k > p \text{ and } \ell = 1 \\ 1 & \text{for all } k = 1 \text{ and } \ell > p \\ \psi(k, \ell) & \text{for all } k \leq p \text{ and } \ell \leq p \\ 0 & \text{else} \end{cases}, \quad (29)$$

while, on the event P_n , the same relations hold for ψ_n . The idea of the proof is to split the analysis into the three regions

$$(k = 1, \ell > p), \quad (k, \ell \leq p) \quad \text{and} \quad (k > p, \ell = 1).$$

The hard work is in the second region. However, since on the event Λ_n all degree are bounded by p , it suffices to analyze individual terms

$$|\psi_n(k, \ell) \mathcal{E}_n(k, \ell) - \psi(k, \ell) \mathcal{E}(k, \ell)|,$$

instead of the full sum

$$\sum_{k, \ell=0}^p |\psi_n(k, \ell) \mathcal{E}_n(k, \ell) - \psi(k, \ell) \mathcal{E}(k, \ell)|.$$

Recall that

$$\Xi_n = \left\{ \sum_{k, \ell=0}^{\infty} |h_n(k, \ell) - h(k, \ell)| \leq K n^{-\delta} \right\}.$$

and let us bound the probability in (28) as follows,

$$\mathbb{P}(\Xi_n^c, \Lambda_n) \leq \mathbb{P}\left(\sum_{k,\ell=1}^{\infty} |\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \quad (30)$$

$$+ \mathbb{P}\left(\sum_{k,\ell=1}^{\infty} \psi(k, \ell) |(\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell))| > \frac{Kn^{-\delta}}{2}, \Lambda_n\right). \quad (31)$$

We will first deal with (30). By (29) and conditioned on Λ_n , we have that $|\psi_n(k, \ell) - \psi(k, \ell)| \neq 0$, only when $k, \ell \leq p$. Hence we get, using the union bound,

$$\begin{aligned} \mathbb{P}\left(\sum_{k,\ell=1}^{\infty} |\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \\ = \mathbb{P}\left(\sum_{k,\ell=1}^p |\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \\ \leq \sum_{k,\ell=1}^p \mathbb{P}\left(|\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > \frac{Kn^{-\delta}}{2p^2}, \Lambda_n\right) \\ \leq O(n^{-\varepsilon+\delta}), \end{aligned}$$

where the last line follows from Lemma 6.2.

Next we consider (31). First we use (29) to bound the term inside the probability as follows

$$\sum_{k,\ell=1}^{\infty} \psi(k, \ell) |(\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell))| \leq \sum_{k=1}^p \sum_{\ell=1}^p \psi(k, \ell) |\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| \quad (32)$$

$$+ \sum_{\ell=p+1}^{\infty} |\mathcal{E}(1, \ell) - \mathcal{E}_n(1, \ell)| \quad (33)$$

$$+ \sum_{k=p+1}^{\infty} |\mathcal{E}(k, 1) - \mathcal{E}_n(k, 1)| \quad (34)$$

We will start by analyzing (33). For this we notice that $\mathcal{E}_n(1, \ell) - \mathcal{E}(1, \ell) = f_n^*(\ell) - f^*(\ell)$, so that

$$\sum_{\ell=p+1}^{\infty} \psi(1, \ell) |\mathcal{E}(1, \ell) - \mathcal{E}_n(1, \ell)| \leq \sum_{\ell=0}^{\infty} |f_n^*(\ell) - f^*(\ell)|.$$

The upper bound for (34) is the same. Therefore, again using the union bound, we have that

$$\begin{aligned} \mathbb{P}\left(\sum_{k,\ell=0}^{\infty} \psi(k, \ell) |(\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell))| > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \\ \leq 2\mathbb{P}\left(|f_n^*(\ell) - f^*(\ell)| > \frac{Kn^{-\delta}}{6}, \Omega_n\right) + \sum_{k,\ell=0}^p \mathbb{P}\left(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > \frac{Kn^{-\delta}}{6p^2}, \Omega_n\right) \\ \leq O(n^{-\varepsilon+\delta}). \end{aligned}$$

Here we used Lemma 6.3 to bound the last probability in the second line.

With this final result we have proven (28) and hence Theorem 3.3. \square

All that is left is to prove the two technical lemmas 6.2 and 6.3. Due to the use of both a minimum and maximum, in the definitions of $\mathcal{E}_n(k, \ell)$ and $\mathcal{E}(k, \ell)$ and the double cases in $\psi_n(k, \ell)$ and $\psi(k, \ell)$, the proofs consists of many case distinctions, where we have to bound each specific case. In order to improve the readability of the proofs we define, for any $k, \ell \geq 0$, the following events

$$\begin{aligned} A_n &= \{1 - F_n^*(k) < F_n^*(\ell)\} \\ B_n &= \{1 - F_n^*(k-1) > F_n^*(\ell-1)\}, \\ I_n &= \{1 - F_n^*(k-1) \leq F_n^*(\ell)\}, \\ J_n &= \{1 - F_n^*(k) \geq F_n^*(\ell-1)\}. \end{aligned}$$

With these definitions we have that $\psi_n(k, \ell) = \mathbb{1}_{\{A_n\}} \mathbb{1}_{\{B_n\}}$. Moreover since $A_n^I \cap B_n^c = \emptyset$ we have that

$$1 - \psi_n(k, \ell) = \mathbb{1}_{\{A_n\}} \mathbb{1}_{\{B_n^c\}} + \mathbb{1}_{\{A_n^c\}} \mathbb{1}_{\{B_n\}}. \quad (35)$$

Where the event A_n and B_n determine the value of $\psi_n(k, \ell)$, so do the events I_n and J_n define the expression for $\mathcal{E}_n(k, \ell)$, as follows:

$$\mathcal{E}_n(k, \ell) = \begin{cases} f_n^*(k) & \text{on the event } I_n \cap J_n \\ 1 - F^*(k-1) - F^*(\ell-1) & \text{on the event } I_n \cap J_n^c \\ F^*(k) + F^*(\ell) - 1 & \text{on the event } I_n^c \cap J_n \\ f_n^*(\ell) & \text{on the event } I_n^c \cap J_n^c. \end{cases} \quad (36)$$

Note that by their definitions,

$$0 \leq \psi_n(k, \ell), \psi(k, \ell), \mathcal{E}_n(k, \ell), \mathcal{E}(k, \ell) \leq 1,$$

for all $k, \ell \geq 0$. In addition we will often use the following result

Lemma 6.4. *Let $k, \ell \geq 0$ be such that $1 - F^*(k) < F^*(\ell)$. Then*

$$\mathbb{P}(1 - F_n^*(k) \geq F_n^*(\ell), \Omega_n) \leq O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$.

If, on the other hand, $1 - F^(k) > F^*(\ell)$, then*

$$\mathbb{P}(1 - F_n^*(k) \leq F_n^*(\ell), \Omega_n) \leq O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$.

Proof. We will prove the first statement, since the proof for the second is similar. First we write

$$\begin{aligned} &\mathbb{P}(1 - F_n^*(k) \geq F_n^*(\ell), \Omega_n) \\ &\mathbb{P}((F^*(k) - F_n^*(k)) + 1 - F^*(k) \geq F^*(\ell) + (F_n^*(\ell) - F^*(\ell)), \Omega_n) \\ &\mathbb{P}((F^*(k) - F_n^*(k)) + (F_n^*(\ell) - F^*(\ell)) \geq F^*(\ell) - 1 + F^*(k), \Omega_n). \end{aligned}$$

Next we use the union bound and Markov's inequality to obtain

$$\begin{aligned} &\mathbb{P}(1 - F_n^*(k) \geq F_n^*(\ell), \Omega_n) \\ &\leq \mathbb{P}(|F^*(k) - F_n^*(k)| \geq F^*(\ell) - 1 + F^*(k), \Omega_n) \\ &\quad + \mathbb{P}(|F^*(\ell) - F_n^*(\ell)| \geq F^*(\ell) - 1 + F^*(k), \Omega_n) \\ &\leq \frac{2\mathbb{E}[\sup_{k \geq 0} |F_n^*(k) - F^*(k)| \mathbb{1}_{\{\Omega_n\}}]}{F^*(\ell) - 1 + F^*(k)} \\ &\leq \mathbb{E}\left[\sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \mathbb{1}_{\{\Omega_n\}}\right] = O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$, where we used $1 - F^*(k) < F^*(\ell)$ for the last equality. \square

Proof of Lemma 6.2. Note that the specific expression of $\psi(k, \ell)$ depends on the ordering between

$$1 - F^*(k) \quad \text{and} \quad F^*(\ell),$$

and

$$1 - F^*(k-1) \quad \text{and} \quad F^*(\ell-1).$$

Therefore, we need to consider all different cases ($<$, $=$, $>$), where we treat equality as a separate case. This gives a total of nine cases. However, there are several combinations that do not need to be considered. For instance, $1 - F^*(k) > F^*(\ell)$ implies that $1 - F^*(k-1) \geq F^*(\ell-1)$. In the end, we are left with the following cases:

- I) $1 - F^*(k) < F^*(\ell)$ and $1 - F^*(k-1) < F^*(\ell-1)$
- II) $1 - F^*(k) = F^*(\ell)$ and $1 - F^*(k-1) < F^*(\ell-1)$
- III) $1 - F^*(k) < F^*(\ell)$ and $1 - F^*(k-1) = F^*(\ell-1)$
- IV) $1 - F^*(k) < F^*(\ell)$ and $1 - F^*(k-1) > F^*(\ell-1)$
- V) $1 - F^*(k) = F^*(\ell)$ and $1 - F^*(k-1) > F^*(\ell-1)$
- VI) $1 - F^*(k) > F^*(\ell)$ and $1 - F^*(k-1) > F^*(\ell-1)$

We will start with the first case.

I) $1 - F^*(k) < F^*(\ell)$ and $1 - F^*(k-1) < F^*(\ell-1)$

First, note that in this case $\psi(k, \ell) = 0$. Moreover, since $F^*(\ell-1) > 1 - F^*(k-1)$, it follows from Lemma 6.4 that

$$\begin{aligned} \mathbb{P}(B_n) &\leq \mathbb{P}(B_n, \Omega_n) + \mathbb{P}(\Omega_n^c) \\ &\leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)). \end{aligned}$$

Hence, since $\psi_n(k, \ell) = 0$ on the event B_n^c , we have

$$\begin{aligned} \mathbb{P}(|\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > Kn^{-\delta}) \\ &= \mathbb{P}(\psi_n(k, \ell) \mathcal{E}_n(k, \ell) > Kn^{-\delta}, B_n^c) + \mathbb{P}(B_n) \\ &\leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)). \end{aligned}$$

II) $1 - F^*(k) = F^*(\ell)$ and $1 - F^*(k-1) < F^*(\ell-1)$

In this case we again have that $\psi(k, \ell) = 0$. In addition

$$1 - F^*(k-1) > 1 - F^*(k) > F^*(\ell),$$

so that, by Lemma 6.4

$$\mathbb{P}(I_n) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

Similarly, using that $1 - F^*(k) > F^*(\ell-1)$, we have

$$\mathbb{P}(J_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

Therefore, using (36) and $1 - F^*(k) = F^*(\ell)$, it follows that

$$\begin{aligned} \mathbb{P}(|\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > Kn^{-\delta}) \\ &= \mathbb{P}(\psi_n(k, \ell) \mathcal{E}_n(k, \ell) > Kn^{-\delta}) \\ &\leq \mathbb{P}(\mathcal{E}_n(k, \ell) > Kn^{-\delta}, I_n^c, J_n) + \mathbb{P}(I_n) + \mathbb{P}(J_n^c) + \mathbb{P}(I_n, J_n^c) \\ &\leq \mathbb{P}(|F_n^*(\ell) + F_n^*(k) - 1| > Kn^{-\delta}) + O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)) \end{aligned}$$

$$\begin{aligned}
&\leq \mathbb{P} \left(|F_n^*(\ell) - F^*(\ell)| > \frac{Kn^{-\delta}}{2} \right) \\
&\quad + \mathbb{P} \left(|F_n^*(k) - F^*(k)| > \frac{Kn^{-\delta}}{2} \right) + O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)) \\
&\leq +O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)),
\end{aligned}$$

where for the fifth line we used that

$$\begin{aligned}
|F_n^*(\ell) + F_n^*(k) - 1| &= |F_n^*(\ell) - F^*(\ell) + (1 - F^*(k)) + F_n^*(k) - 1| \\
&\leq |F_n^*(\ell) - F^*(\ell)| + |F_n^*(k) - F^*(k)|,
\end{aligned}$$

since $1 - F^*(k) = F^*(\ell)$.

Case III) and V) can be dealt with using arguments similar to case II), while case VI) is similar to I). Therefore, there is only one case left.

IV) $1 - F^*(k) < F^*(\ell)$ and $1 - F^*(k - 1) > F^*(\ell - 1)$

We first note that, since $1 - F^*(k) - F^*(\ell) > 0$,

$$\mathbb{P}(A_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)),$$

by Lemma 6.4, and similarly

$$\mathbb{P}(B_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

Since for this case $\psi(k, \ell) = 1$, we have,

$$\begin{aligned}
&\mathbb{P}(|\psi_n(k, \ell) - \psi(k, \ell)| \mathcal{E}_n(k, \ell) > Kn^{-\delta}) \\
&= \mathbb{P}((1 - \psi_n(k, \ell))\mathcal{E}_n(k, \ell) > Kn^{-\delta}) \\
&\leq \mathbb{P}((1 - \psi_n(k, \ell))\mathcal{E}_n(k, \ell) > Kn^{-\delta}, A_n, B_n) \\
&\leq \mathbb{P}(A_n^c) + \mathbb{P}(B_n^c) + \mathbb{P}(A_n^c, B_n^c) \\
&\leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)),
\end{aligned}$$

where we used (35) for the third line. □

Proof of Lemma 6.3. Similar to the proof of Lemma 6.2 we will have to consider different cases. Here these are with respect to the different relations between

$$1 - F^*(k - 1) \quad \text{and} \quad F^*(\ell),$$

and

$$1 - F^*(k) \quad \text{and} \quad F^*(\ell - 1),$$

which determine the expression for $\mathcal{E}(k, \ell)$. To analyze each case we will also need to distinguish between the different expression of $\mathcal{E}_n(k, \ell)$, which are determined by the events I_n and J_n .

We will consider the three cases where $1 - F^*(k) > F^*(\ell - 1)$. The other six cases can be dealt with using similar arguments. First note that by Lemma 6.4

$$\mathbb{P}(J_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

I) $1 - F^*(k - 1) < F^*(\ell)$ and $1 - F^*(k) > F^*(\ell - 1)$

Similar to $\mathbb{P}(J_n^c)$, it follows from Lemma 6.4 that

$$\mathbb{P}(I_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

Therefore, by conditioning on the different combinations of I_n and J_n , we get

$$\mathbb{P}(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > Kn^{-\delta})$$

$$\begin{aligned}
&\leq \mathbb{P}(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > Kn^{-\delta}, I_n, J_n) + \mathbb{P}(J_n^c) + \mathbb{P}(I_n^c) + \mathbb{P}(J_n^c, I_n^c) \\
&\leq \mathbb{P}(|f^*(k) - f_n^*(k)| > Kn^{-\delta}, \Omega_n) + O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)) \\
&\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).
\end{aligned}$$

II) $1 - F^*(k-1) = F^*(\ell)$ and $1 - F^*(k) > F^*(\ell-1)$

Since $f^*(k) = F^*(k) - F^*(k-1)$,

$$\begin{aligned}
|F_n^*(k) + F_n^*(\ell) - 1 - f^*(k)| &= |F_n^*(k) - F^*(k) + F_n^*(\ell) - 1 + F^*(k-1)| \\
&\leq |F_n^*(k) - F^*(k)| + |F_n^*(\ell) - F^*(\ell)|,
\end{aligned}$$

from which it follows that

$$\begin{aligned}
&\mathbb{P}(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > Kn^{-\delta}, I_n^c, J_n) \\
&\leq \mathbb{P}(|F_n^*(k) + F_n^*(\ell) - 1 - f^*(k)| > Kn^{-\delta}) \\
&\leq \mathbb{P}\left(|F_n^*(k) - F^*(k)| > \frac{Kn^{-\delta}}{2}\right) + \mathbb{P}\left(|F_n^*(\ell) - F^*(\ell)| > \frac{Kn^{-\delta}}{2}\right) \\
&\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
&\mathbb{P}(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > Kn^{-\delta}) \\
&\leq \mathbb{P}(|f^*(k) - f_n^*(k)| > Kn^{-\delta}) \\
&\quad + \mathbb{P}(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > Kn^{-\delta}, I_n^c, J_n) + 2\mathbb{P}(J_n^c) \\
&\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).
\end{aligned}$$

III) $1 - F^*(k-1) > F^*(\ell)$ and $1 - F^*(k) > F^*(\ell-1)$

First we notice that in this case $\mathcal{E}(k, \ell) = F^*(\ell) + F^*(k) - 1$. Next, using Lemma 6.4, we have

$$\mathbb{P}(I_n) \leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n)).$$

Therefore it follows that

$$\begin{aligned}
&\mathbb{P}(|\mathcal{E}(k, \ell) - \mathcal{E}_n(k, \ell)| > Kn^{-\delta}) \\
&\leq \mathbb{P}(|F^*(\ell) + F^*(k) - 1 - \mathcal{E}_n(k, \ell)| > Kn^{-\delta}, I_n^c, J_n) \\
&\quad + \mathbb{P}(I_n) + 2\mathbb{P}(J_n^c) \\
&\leq \mathbb{P}\left(|F_n^*(k) - F^*(k)| > \frac{Kn^{-\delta}}{2}\right) + \mathbb{P}\left(|F_n^*(\ell) - F^*(\ell)| > \frac{Kn^{-\delta}}{2}\right) \\
&\quad + \mathbb{P}(I_n) + 2\mathbb{P}(J_n^c) \\
&\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).
\end{aligned}$$

□

6.5 Main results

Here we will give the proofs of our two main results. We start with a useful result which we need to prove Theorem 2.1.

Proposition 6.5. *Let $G_n \in \mathcal{G}_{\eta, \varepsilon}(f, f^*)$ and let X, Y be random variables with joint distribution h as defined in (10). Then, for any $0 < \delta < \varepsilon$ and $K > 0$,*

$$\mathbb{P}(|\tilde{\rho}(G_n) - \rho(D_*, D^*)| > n^{-\delta}) = O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$.

Proof. First we write

$$\begin{aligned}
\left| \tilde{\rho}(\hat{G}_n) - \rho(X, Y) \right| &\leq 3 \left| \sum_{k, \ell=0}^{\infty} \mathcal{F}_n^*(k) \mathcal{F}_n^*(\ell) h_n(k, \ell) - \mathcal{F}^*(k) \mathcal{F}^*(\ell) h(k, \ell) \right| \\
&\leq 3 \left| \sum_{k, \ell=0}^{\infty} \mathcal{F}_n^*(k) \mathcal{F}_n^*(\ell) (h_n(k, \ell) - h(k, \ell)) \right| \\
&\quad + 3 \sum_{k, \ell=0}^{\infty} |\mathcal{F}_n^*(k) \mathcal{F}_n^*(\ell) - \mathcal{F}^*(k) \mathcal{F}^*(\ell)| h_n(k, \ell) \\
&\leq 12 \sum_{k, \ell=0}^{\infty} |h_n(k, \ell) - h(k, \ell)| + 24 \sup_k |F_n^*(k) - F^*(k)|. \tag{37}
\end{aligned}$$

For the last inequality, we used

$$\begin{aligned}
&\sum_{k, \ell=0}^{\infty} |\mathcal{F}_n^*(k) \mathcal{F}_n^*(\ell) - \mathcal{F}^*(k) \mathcal{F}^*(\ell)| h_n(k, \ell) \\
&\leq \sup_{k, \ell} |\mathcal{F}_n^*(k) \mathcal{F}_n^*(\ell) - \mathcal{F}^*(k) \mathcal{F}^*(\ell)| \\
&\leq \sup_{k, \ell} |\mathcal{F}_n^*(k) - \mathcal{F}^*(k)| \mathcal{F}_n^*(\ell) + \sup_{k, \ell} |\mathcal{F}_n^*(\ell) - \mathcal{F}^*(\ell)| \mathcal{F}^*(k) \\
&\leq 4 \sup_k |\mathcal{F}_n^*(k) - \mathcal{F}^*(k)| \leq 8 \sup_k |F_n^*(k) - F^*(k)|.
\end{aligned}$$

Note that by Theorem 3.3

$$\mathbb{P} \left(12 \sum_{k, \ell=0}^{\infty} |h_n(k, \ell) - h(k, \ell)| > \frac{n^{-\delta}}{2} \right) = O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).$$

Moreover, on the event Ω_n ,

$$\sup_{k \geq 0} |F_n^*(k) - F^*(k)| \leq \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \leq n^{-\varepsilon}.$$

Hence, it follows from (37) and Markov's inequality that

$$\begin{aligned}
\mathbb{P} \left(\left| \rho(\tilde{G}_n) - \rho(X, Y) \right| > n^{-\delta} \right) &\leq \mathbb{P} \left(12 \sum_{k, \ell=0}^{\infty} |h_n(k, \ell) - h(k, \ell)| > \frac{n^{-\delta}}{2} \right) \\
&\quad + \mathbb{P} \left(24 \sup_k |F_n^*(k) - F^*(k)| > \frac{n^{-\delta}}{2}, \Omega_n \right) + O(\mathbb{P}(\Omega_n^c)) \\
&\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)) + 48n^\delta \mathbb{E} \left[\sup_k |F_n^*(k) - F^*(k)| \mathbb{1}_{\{\Omega_n\}} \right] \\
&\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).
\end{aligned}$$

□

We are now ready to give the proof of Theorem 2.1.

Proof of Theorem 2.1. Consider a graph $G_n \in \mathcal{G}_{n, \varepsilon}(f, f^*)$, denote its degree sequence by \mathbf{D}_n , let $\tilde{G}_n = \text{DGA}(\mathbf{D}_n)$ and recall that $\kappa = (\varepsilon + \delta)/2$. Then, since $\delta < \kappa < \varepsilon$, it follows from Proposition 6.5 that

$$\mathbb{P} \left(\left| \tilde{\rho}(\tilde{G}_n) - \rho(D_*, D^*) \right| > K n^{-\delta} \right) \leq O(n^{-\varepsilon+\kappa} + \mathbb{P}(\Omega_n^c)),$$

which proves the second statement of the theorem.

For the first statement, note that by Theorem 3.4

$$\sum_{i \rightarrow j \in G_n} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j) \geq \sum_{i \rightarrow j \in \tilde{G}_n} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j)$$

so that

$$\tilde{\rho}(G_n) \geq \tilde{\rho}(\tilde{G}_n).$$

Therefore we have, as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(\tilde{\rho}(G_n) < \rho(D_*, D^*) - Kn^{-\delta}) &\leq \mathbb{P}(\tilde{\rho}(\tilde{G}_n) < \rho(D_*, D^*) - n^{-\delta}) \\ &\leq \mathbb{P}(|\tilde{\rho}(\tilde{G}_n) - \rho(D_*, D^*)| > Kn^{-\delta}) \\ &\leq O(n^{-\varepsilon - \kappa} + \mathbb{P}(\Omega_n^c)), \end{aligned}$$

which proves the first statement of the theorem. \square

We now move on to Theorem 2.2. We will follow the strategy described in Section 4, that is we will use the delta transformation to construct a degree distribution f_ρ for which $f_\rho^*(1)$ is large enough.

First observe that (22) together with Proposition 6.5 imply 4.1.

Proof of Theorem 2.2. Let δ be such that $9(\xi/2)^2 - 6(\xi/2)^3 - 3 = \rho + \epsilon$, for some $\epsilon > 0$, and denote by f^* the size-biased distribution of f . Now take f_δ^* to be the δ -transform of f^* and set

$$\mu_\rho = \left(\mu(1 - F(K_\delta)) + \sum_{t=1}^{K_\delta} \frac{f_\delta^*(t)}{t} \right)^{-1},$$

where K_δ was defined as the smallest integer such that $F^*(K_\delta) > \delta$. Now we define the function f_ρ by:

$$f_\rho(0) = \frac{\mu_\rho f_\delta^*(1)}{2} = f_\rho(1) = \frac{\mu_\rho f_\delta^*(1)}{2} \quad \text{and} \quad f_\rho(t) = \frac{\mu_\rho f_\delta^*(t)}{t} \quad \text{for } k \geq 2.$$

Then, since by construction $f_\rho^*(t) = f^*(t)$ for all $t > K_\delta$, it follows that

$$\begin{aligned} \sum_{t=0}^{\infty} f_\rho(t) &= \sum_{t=1}^{\infty} \frac{\mu_\rho f_\rho^*(t)}{t} \\ &= \mu_\rho \left(\sum_{t=1}^{K_\delta} \frac{f_\rho^*(t)}{t} + \sum_{t=K_\delta+1}^{\infty} \frac{f^*(t)}{t} \right) \\ &= \mu_\rho \left(\sum_{t=1}^{K_\delta} \frac{f_\delta^*(t)}{t} + \mu(1 - F(K_\delta)) \right) = 1, \end{aligned}$$

so that f_ρ defines a probability density. Moreover, since for all $k > K_\delta$

$$1 - F_\rho(k) = \sum_{t=k+1}^{\infty} f_\rho(t) = \mu_\rho \sum_{t=k+1}^{\infty} \frac{f_\delta^*(t)}{t} = \mu_\rho \sum_{t=k+1}^{\infty} \frac{f^*(t)}{t} = \frac{\mu_\rho}{\mu} \sum_{t=k+1}^{\infty} f(t),$$

it follows that $\sum_{k=0}^{\infty} t^{1+\eta} f_\rho(t) < \infty$ and

$$\lim_{k \rightarrow \infty} \frac{1 - F_\rho(k)}{1 - F(k)} = \frac{\mu_\rho}{\mu}.$$

Now let D have probability density f_ρ , and hence size-biased density $f_\rho^*(t) = tf_\rho(t)/\mu_\rho$, and let \mathbf{D}_n be generated by the IID algorithm, by sampling from D . Then, by Lemma 3.1, $\mathbf{D}_n \in \mathcal{D}_{\eta,\varepsilon}(f_\rho, f_\rho^*)$ and since by construction of f_ρ we have that $f_\rho^*(1) = \delta/2$, it follows that

$$9f_\rho^*(1)^2 - 6f_\rho^*(1)^3 - 3 = \rho + \epsilon.$$

Hence, if G_n is a graph with degree sequence \mathbf{D}_n , we have, by taking $\delta = \min(\varepsilon, 1/2)/2$ in Proposition 4.1, that as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(\rho(G_n) > \rho) &\geq \lim_{n \rightarrow \infty} \mathbb{P}(\rho(G_n) > \rho + \epsilon - n^{-\delta}) \\ &= \mathbb{P}(\rho(G_n) > 9f_\rho^*(1)^2 - 6f_\rho^*(1)^3 - 3 - n^{-\delta}) \\ &\geq 1 - O\left(n^{-\varepsilon+3\kappa/4} + \mathbb{P}(\Omega_n^c)\right). \end{aligned}$$

□

References

- [1] David L Alderson and Lun Li. Diversity of graphs with highly variable connectivity. *Physical Review E*, 75(4):046102, 2007.
- [2] Kevin E Bassler, Charo I Del Genio, Péter L Erdős, István Miklós, and Zoltán Toroczkai. Exact sampling of graphs with prescribed degree correlations. *New Journal of Physics*, 17(8):083052, 2015.
- [3] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.
- [4] Ningyuan Chen and Mariana Olvera-Cravioto. Efficient simulation for branching linear recursions. *arXiv preprint arXiv:1503.09150*, 2015.
- [5] Philippe Deprez and Mario V Wüthrich. Construction of directed assortative configuration graphs. *arXiv preprint arXiv:1510.00575*, 2015.
- [6] TR Hurd. The construction and properties of assortative configuration graphs. *arXiv preprint arXiv:1512.03084*, 2015.
- [7] Ravi Kannan, Prasad Tetali, and Santosh Vempala. Simple markov-chain algorithms for generating bipartite graphs and tournaments. *Random Structures and Algorithms*, 14(4):293–308, 1999.
- [8] Sergei Maslov and Kim Sneppen. Specificity and stability in topology of protein networks. *Science*, 296(5569):910–913, 2002.
- [9] Jörg Menche, Angelo Valleriani, and Reinhard Lipowsky. Asymptotic properties of degree-correlated scale-free networks. *Physical Review E*, 81(4):046103, 2010.
- [10] Mhamed Mesfioui and Abdelouahid Tajar. On the properties of some nonparametric concordance measures in the discrete case. *Nonparametric Statistics*, 17(5):541–554, 2005.
- [11] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(2-3):161–180, 1995.
- [12] Mark EJ Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.
- [13] Mark EJ Newman, Steven H Strogatz, and Duncan J Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2):026118, 2001.

- [14] Isabelle Stanton and Ali Pinar. Constructing and sampling graphs with a prescribed joint degree distribution. *Journal of Experimental Algorithmics (JEA)*, 17:3–5, 2012.
- [15] Remco van der Hofstad. Random graphs and complex networks. *Lecture notes in prep*, 2014.
- [16] Remco van der Hofstad and Nelly Litvak. Degree-degree dependencies in random graphs with heavy-tailed degrees. *Internet mathematics*, 10(3-4):287–334, 2014.
- [17] Pim van der Hoorn and Nelly Litvak. Convergence of rank based degree-degree correlations in random directed networks. *Moscow Journal of Combinatorics and Number Theory*, 4(4):45–83, 2014.
- [18] Pim van der Hoorn and Nelly Litvak. Degree-degree dependencies in directed networks with heavy-tailed degrees. *Internet Mathematics*, 11(2):155–179, 2015.
- [19] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.
- [20] R Xulvi-Brunet and IM Sokolov. Reshuffling scale-free networks: From random to assortative. *Physical Review E*, 70(6):066102, 2004.
- [21] Dan Yang, Liming Pan, and Tao Zhou. Lower bound of assortativity coefficient in scale-free networks. *arXiv preprint arXiv:1602.04350*, 2016.